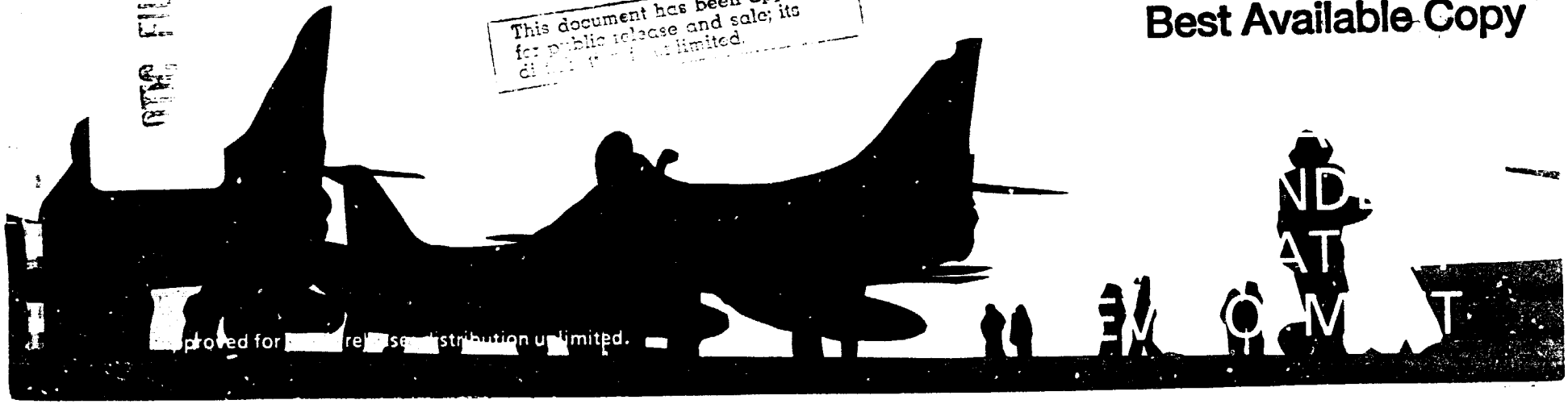
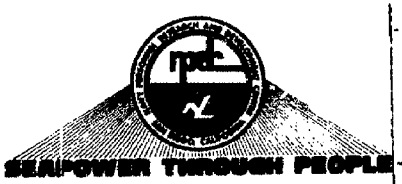
Navy Personnel Research and Development Center

San Diego, CA 92152-6800 SR 87-1 January 1987

ADA 176944





Navy Personnel Research and Development Center
San Diego, California 92152-6800

FOR WORD

The technical directors of Navy Laboratories receive Independent Research (IR) and Independent Exploratory Development (IED) funding annually to support innovative, promising work without going through the formal approval procedures required under normal funding authorization. This funding enables the researchers selected for the program to devote a portion of their time to basic research that offers potential benefits for specific military problem are .

Research at the Navy Personnel Research and Development Center addresses the Navy's needs for enhancing system and personnel performance through the integration of people and technology. Resources provided for the IR/IED program have been used to develop a variety of research methods, models, and techniques within the areas of training, manpower utilization, organizational productivity, and human factors engineering of naval weapon systems and platforms.

The IR program has been active at this Center since 1973 and is funded under Program Element (PE) 61152N. The IED program was initiated in 1976 and is funded by PE 62766N.

This report is submitted to fulfill the requirement for an annual IR/IED report. It provides summaries of selected FY86 projects, the program funding profile, and a list of FY86 publications and presentations resulting from IR/IED efforts.

B. E. Bacon
Captain, U.S. Navy
Commanding Officer

J. W. Tweeddale
Technical Director

NPRDC Special Report 87-1

January 1987

FY86 INDEPENDENT RESEARCH/ INDEPENDENT EXPLORATORY DEVELOPMENT

Alice M. Crawford
Margen H. Metcalfe
(Editors)

Reviewed by
J. W. Tweeddale

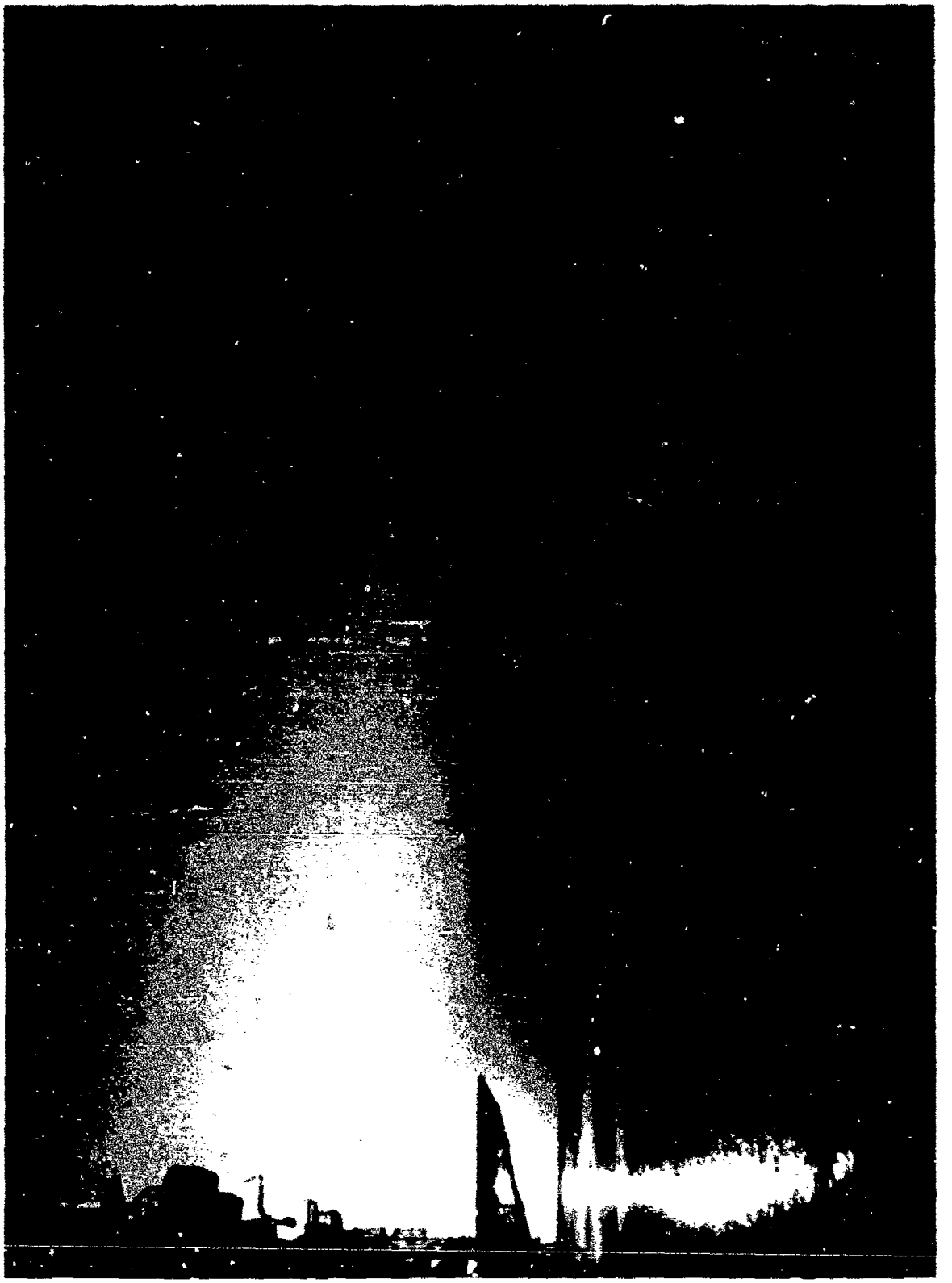
Released by
B. E. Bacon
Captain, U. S. Navy
Commanding Officer

Navy Personnel Research and Development Center
San Diego, California 92152-6800

CONTENTS

CONTENTS

CONTENTS



NAVY PERSONNEL RESEARCH AND DEVELOPMENT CENTER
SAN DIEGO, CALIFORNIA
npdc

SEAPOWER THROUGH PEOPLE

J. B. (Brad) Sympson

J. B. (Brad) Sympson is one of a handful of individuals in this country who has developed a high level of expertise in Item Response Theory (IRT). IRT is a new approach to mental testing that is revolutionizing the way tests are designed, administered, and scored. After completing advanced graduate work in psychometrics (the theory of mental testing) at the University of Minnesota, Brad spent two years at Educational Testing Service working with Dr. Frederic M. Lord, one of the initial developers of IRT. Brad came to the Navy Personnel Research and Development Center (NAVPERSRANDCEN) in 1981 to work on the Joint Services Computerized Adaptive Testing (CAT) Project. Since coming to NAVPERS-RANDCEN, Brad has developed several new statistical procedures and IRT models for use in CAT. He also serves as a consultant to other researchers on the CAT Project.

Models for Calibrating Multiple-choice Items

J. B. (Brad) Sympson

Independent Research

Dichotomous (right/wrong) scoring of multiple-choice test questions does not distinguish among various wrong answers chosen by examinees. Wrong answers can supply valuable information about examinee ability. A new item-response model and a polychotomous item-scoring procedure, developed through this research, have several benefits for selecting and classifying military personnel.

BACKGROUND

Mental Testing

For more than 40 years, selection and classification of enlisted military personnel has depended on objective tests of mental ability. These tests are especially appropriate for selecting and classifying individuals who lack specialized training or experience and must undergo formal training in preparation for their job assignment.

The benefits of mental ability tests go far beyond identifying individuals who are likely to perform well during their initial enlistment. Since the armed services promote from within, the quality of personnel accepted for initial entry ultimately determines the quality of personnel available for the upper enlisted ranks. Thus, both short-term and long-term outcomes rely heavily on decisions made prior to initial enlistment. Mental ability tests help increase the likelihood that these decisions will be good ones.

Multiple-choice Questions

Paper-and-pencil tests such as the Armed Services Vocational Aptitude Battery (ASVAB; Department of Defense, 1984) typically consist of multiple-choice questions. Although computerized testing will probably reduce the use of paper-and-pencil tests, multiple-choice questions will continue to be widely used. Even when an examinee is asked to enter a "free response" on a computer (which requires the examinee to recall, rather than recognize, the correct answer), the computer must assign the response to one of several predefined, mutually exclusive categories. Thus, even test questions that are not presented in a multiple-choice format will usually be scored as if they were multiple-choice questions.

PROBLEM

In current applications of multiple-choice questions to ability testing (e.g., the ASVAB and computerized testing), examinee responses are scored as either correct

or incorrect. This dichotomous item-scoring procedure does not distinguish among the various incorrect answers that examinees select. Information about examinee ability that could be extracted from wrong answers is lost.

Also, currently-used dichotomous item-response models fail to "fit" a portion of the multiple-choice questions that are written by test developers. If an item-response model is to be used, items that do not fit the model must be set aside. This reduces the number of items that are available for use during testing.

OBJECTIVE

The objective of this research was to develop new psychometric (psychological measurement) procedures that will extract additional information about an examinee's level of ability from the examinee's wrong answers. This will be accomplished through the development of polychotomous item-response models, new testing strategies based on these models, and methods of scoring tests that extract information about examinee ability from all item responses--incorrect as well as correct.

PROGRESS

Several polychotomous item-response models have been developed and tried on actual test data (Sympson, 1981, 1983, 1986a). The most promising of these models is being evaluated (Sympson, 1986b). A new family of statistical distribution functions and a computer program that fits this distribution function to sets of test scores have been developed (Sympson & France, 1984). A computer program that uses polychotomous scoring during item analysis has also been developed and applied to several sets of test data (Sympson, 1984).

EXAMPLE

A 25-item test of arithmetic reasoning ability taken by 1300 Marine Corps recruits was analyzed using the item analysis program developed in this research. One of the questions analyzed was the following:

> *A key rack has 8 rows of hooks. Each row has 6 hooks. If 25 percent of the hooks have keys on them, how many hooks are empty?*
>
> *(a) 12*
> *(b) 16*
> *(c) 32*
> *(d) 36* *(correct answer)*
>
> *Solving this problem requires three steps:*
>
> *(1) 8 x 6 = 48 hooks*
> *(2) 48 x .25 = 12 hooks with keys.*
> *(3) 48 - 12 = 36 empty hooks.*

Of the Marine recruits tested, 33 percent selected option "d," the correct answer. Another 33 percent selected option "a." The remaining 34 percent selected either "b" or "c." Apparently, individuals who selected option "a" completed the first two steps in the solution and then stopped. Although option "a" is incorrect, choosing this option clearly indicates a higher level of ability than choosing either "b" or "c," which are unrelated to the sequence of steps required to solve the problem.

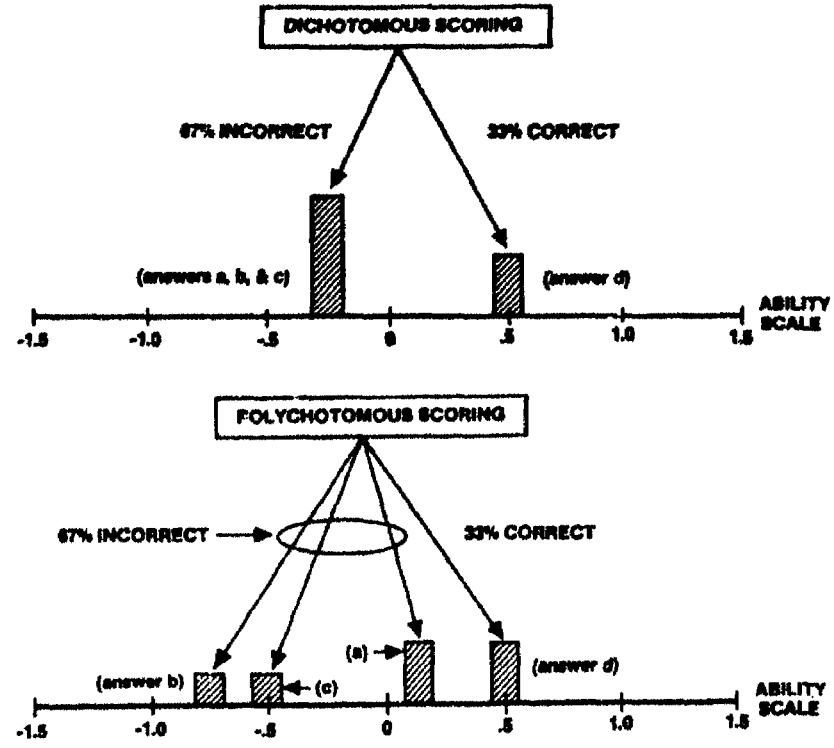


Figure 1. Models for calibrating multiple-choice items.

The upper portion of Figure 1 shows the result of scoring this four-choice item dichotomously. Examinees who selected the correct answer were assigned a positive ability-level estimate, while examinees who answered incorrectly were assigned a negative ability-level estimate.

The lower portion of Figure 1 shows the result of scoring this same item polychotomously. Examinees who answered correctly were assigned the same ability estimate as before, but examinees who answered incorrectly were assigned three different ability estimates, depending on which incorrect answer they

chose. In particular, examinees who selected response-option "a" received an ability estimate that is positive, but lower than the one assigned to examinees who answered correctly. Sorting people who answer incorrectly into different groups provides additional information about their mental ability and serves to increase test reliability.

This example demonstrates how additional information about an examinee's ability level can be extracted by considering which incorrect answers have been selected. It also shows that treating all wrong answers as equivalent can be unfair to those examinees who have given "partially correct" answers.

The importance of option "a" in this item was discovered using the psychometric procedures developed in this research. These procedures are based on statistical analyses of examinee item responses. They do not require one to read each question in an attempt to discover the relationship between ability and wrong answers.

BENEFITS

1. Empirical results (Sympson, 1986b) indicate that polychotomous item-response models do provide additional information about examinee ability. Application of polychotomous models will allow us to shorten mental ability tests by about 15-20 percent, without sacrificing test reliability.

2. The best polychotomous model developed in this research has "fit" every test item to which it was applied. Thus, if this model is implemented, more of the test questions that are written can be used.

3. Our procedures allow test developers to identify test questions and response alternatives that are especially good or especially poor indicators of mental ability and aid in determining the nature of the psychological processes that underlie examinee responses.

PLANS

During FY87, our most promising polychotomous item-response model will be applied to a wider variety of test questions. We will also document the various computer programs that have been developed and will report our research findings in the technical literature.

All of these benefits will serve to improve the measurement of mental ability in military selection and classification tests.

REFERENCES

Department of Defense (1984). Test Manual for the Armed Services Vocational Aptitude Battery (DoD 1304.12AA). North Chicago, IL: U.S. Military Entrance Processing Command.

Sympson, J. B. (October 1981). A nominal model for IRT item calibration. Talk given at the Office of Naval Research Conference on Model-based Measurement, Millington,

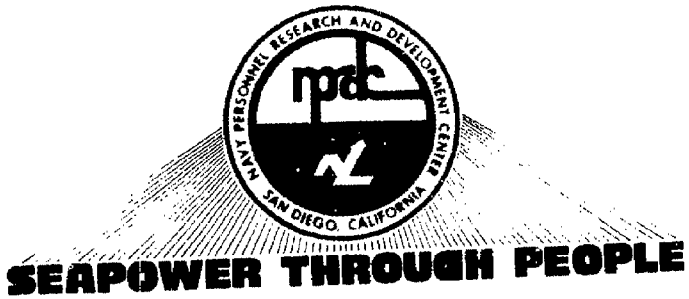
Sympson, J. B. (June 1983). A new item response theory model for calibrating multiple-choice items. Paper presented at the meeting of the Psychometric Society, Los Angeles, CA.

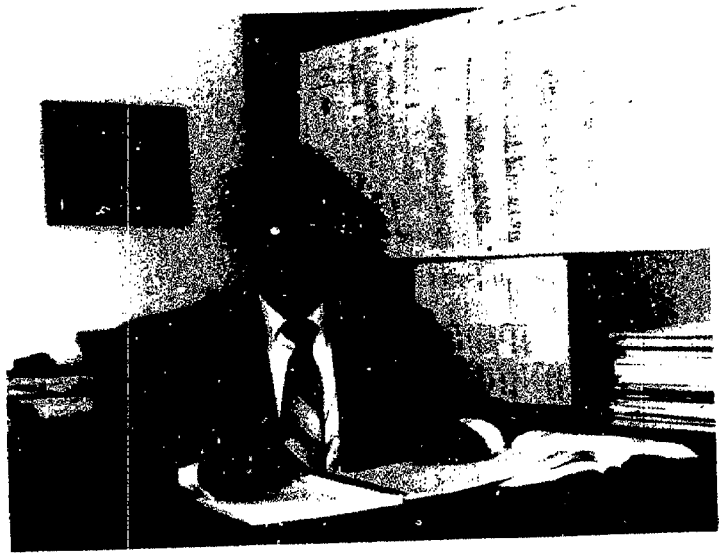
Sympson, J. B. (October 1984). Principal component analysis of polychotomous item responses. Talk given at the Office of Naval Research Conference on Model-based Psychological Measurement, Princeton, NJ.

Sympson, J. B. (April 1986a). Some item response functions obtained in polychotomous item analysis. Talk given at the Office of Naval Research Conference on Model-based Psychological Measurement, Gatlinburg, TN.

Sympson, J. B. (August 1986b). Extracting information from wrong answers in computerized adaptive testing. In B. F. Green (Chair), New developments in computerized adaptive testing. Symposium conducted at the meeting of the American Psychological Association, Washington, DC.

Sympson, J. B., & France, S. L. (October 1984). A family of distribution functions. Paper presented at the Office of Naval Research Conference on Model-based Psychological Measurement, Princeton, NJ.





Timothy T. Liang

Dr. Timothy Liang holds a B.A. degree and M.S. degree in economics, and a Ph.D. in Operations Research. He spent many years working with the private industry and the Hawaii State government. His major research interest was in the area of developing operations research and econometric models to improve manpower planning and resources utilization. Dr. Liang came to the Navy Personnel Research and Development Center in 1980. His major interest is developing mathematical models to solve the Navy's practical problems. His accomplishment is the application of network optimization techniques to the personnel assignment problem and policy research.

Approaches to Multiple Objective Assignment

Timothy T. Liang **Independent Research**

The Navy enlisted personnel assignment process matches over 200,000 people to jobs each year. More than just an administrative necessity, the process is used to continually reassess and realign human resources for better operational readiness. Earlier IR/IED efforts developed a large-scale computer model that helps assign nonrated personnel quickly and according to the latest priority of Navy policies. That version was recently upgraded for use with selected ratings. The current IR/IED effort makes the model even more powerful by incorporating training school quotas, an additional large-scale constraint on assignment. In a typical test case, the model improved the utilization of existing personnel skills by 10 percent, compared to a simulation of the manual matching process. The results indicate that the model makes maximum use of current skills and reduces the need for additional training resources.

SUMMARY

Previous IR/IED efforts formulated the Navy's assignment problem as a pure network, multiple-objective transshipment model. This model has been accepted for implementation by the Enlisted Personnel Management Center, New Orleans for nonrated personnel, and by the Naval Military Personnel Command for a few ratings in the administrative/deck/supply communities. The current IR effort formulates a more complex assignment problem as a network with "side" (or additional aggregate) constraints model. With this formulation, assignment problems for ratings that require advanced technical training can be solved. By optimizing enroute training decisions in the assignment process, use of this formulation will lead to improved skill and training resources utilization.

BACKGROUND

The Navy needs new analytic methods for evaluating and executing assignment policies. These methods must be powerful enough to handle large scale optimization problems with multiple criteria and integer variables in complex relationships. There are three technical issues to be addressed: aggregation/disaggregation in distribution planning and assignment execution, large-scale integer problems with "side constraints," and the ordering/weighting of multiple criteria.

In Navy assignment, aggregate plans to budget permanent change of station (PCS) moves and allocate class seats in advanced training schools impose constraints on assignments. Assignment procedures, using the aggregate plans as global constraints, focus on

sequential assignment decisions as to which jobs should be filled, and how should people be matched to jobs--by what criteria? The capability for disaggregation, the systematic process of refining the aggregate plans to more detailed plans, currently does not exist.

Previous IR/IED efforts formulated the Navy's assignment problem as a multiple objective transshipment model without the aggregate constraints of school quotas. The model uses a highly efficient capacitated transshipment code. This model has been accepted for implementation by the Enlisted Personnel Management Center (EMPAC), New Orleans for nonrated personnel (seaman, fireman, and airman apprentice) and by the Naval Military Personnel Command for a few ratings in the administrative/deck/supply communities.

PROBLEM

The pure network approach for the assignment problem is adequate for ratings that do not require advanced technical training or multiple skills. However, most ratings require advanced training to meet job requirements. If a person has one of the skills to match one of the job's skill requirements, providing that the person meets all other requirements, the person is considered eligible for assignment. If a person does not meet any one of the skill requirements of a particular job, the person may be sent to a technical school enroute. The decision to send a person to school depends on, among other things, the number of available class seats and other school constraints. These constraints destroy the network structure of the problem. Since the Navy sends about 150,000 people a year to schools for over 1,000 types of technical training, a capability to handle these aggregate constraints needed to be developed.

OBJECTIVE

The objective of this research was to advance the technology to solve the large-scale multiple objective assignment problem with aggregate-level constraints.

APPROACH

The integer programming and network approaches were both investigated. The integer programming approach is capable of handling aggregate constraints, but it is not efficient for large-scale problems. The only feasible approach was to formulate aggregate school constraints within the network problem.

A highly technical occupation in the Navy, the aviation electronics technician, was selected to formulate the network with an aggregate school constraints model. The test was based on a problem of 200 people, 230 jobs, and 43 class seats. The objective was to minimize the total cost of training needed to match people to jobs.

RESULTS

The model was able to assign all 200 people to jobs: 169 people matched jobs directly using their existing skills, while 31 people were scheduled for school training to learn the new skills required. For the same case, a simulated manual process assigned only 181 people: 164 people were assigned directly; 17 were assigned to schools enroute; and 19 more people need school training to meet job requirements representing a shortage of 19 seats for various training. It represents an optimal way for utilizing current skills and training resources.

PLANS

The research will improve the current assignment decision process for advanced technical ratings. The transition to a 6.3 project (Enlisted Personnel Assignment System) is underway where applications to more ratings will be developed. In addition, the capability of assigning people to schools for enroute training opens opportunities for more research to improve policy planning for technical training.

EXPECTED PAYOFF

The Navy spends hundreds of millions of dollars a year for advanced technical training. Improvement in skill and training resource utilization will not only improve personnel readiness but also generate training cost efficiencies.

P-A.Federico

Pat-Anthony Federico is a research psychologist at the Navy Personnel Research and Development Center, San Diego. He earned his B.A. (cum laude) from the University of St. Thomas in 1965 with a double major in mathematics and philosophy and a minor in physics. He was awarded his Ph.D. in 1969 from Tulane University in general experimental psychology. He has research interests in individual differences in cognitive processing, learning, and performance; and computer-based instruction and performance assessment. He was elected and served as Executive Director, President, and Secretary-Treasurer of the Human Factors Society, San Diego Chapter. He is also a member of the Cognitive Science Society, Psychonomic Society, American Psychological Association, and the American Educational Research Association. He is a member of the editorial advisory review board for the Journal of Educational Psychology. He has authored or edited over 75 scientific contributions including books, chapters, journal articles, professional papers, and technical reports. His biography is included in Who's Who in Frontier Science and Technology, Who's Who in the West, Who's Who in California, Men of Achievement (Cambridge, England), as well as others.

Performance on Computer-based Cognitive Tasks [1]

P-A.Federico

Independent Research

The objective of this research is to study the process of stabilization during skill acquisition in performance on a computer-based simulation of a complex cognitive task. Software has been developed to construct dynamic graphic simulations with which naval officers can allocate, deploy, and manage fighter and support aircraft to defend carrier-based naval task forces against Soviet naval air bombers. An infinite number of tactical scenarios can be generated. Individuals will be assessed on how well they manage the simulated outer air battle, and the data will be used to create a theoretical framework as well as build the basis of an expert system for a computer-based intelligent tactician.

BACKGROUND

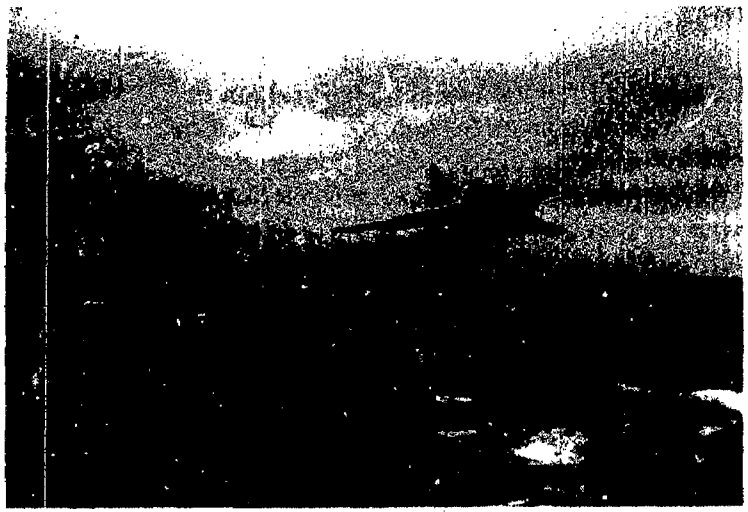
Individuals vary in their rates and manners of skill acquisition, especially in the beginning of practice, and they reach terminal performance plateaus differentially. Early performance requires high conscious control (i.e., it is slow, sequential, effortful, limited, and directed), whereas late performance tends to be automatic (i.e., it is fast, parallel, effortless, and less limited by attentional focus). Practice during the early stages results in dramatic changes in behavior (e.g., decreasing performance variability, minimizing response time). With practice, rate of improvement diminishes and becomes more uniform across individuals (i.e., performance stabilizes). For some tasks, performance does not seem to get any better or worse, and curves that reflect the rate of skill acquisition of individuals appear to be parallel (Ackerman & Schneider, 1984; Jones, 1984; Schneider, 1984). Individual variability among learners affects modes and speeds of skill acquisition: distinct experiences, cognitive models,, aptitudes, and motivations can influence early and late performance differentially.

Much of the earlier research on which the above statements are based was done with psychomotor tasks. Considerably less is known about complex tasks that are primarily cognitive in nature.

PROBLEM

Because many factors affect the nature and time course of skill acquisition, beginning performance on complicated tasks is usually not a good estimate of

[1] Cost shared with Center 6.2 funding ($40K from PE 62763N).



terminal performance. Since initially intricate performance does not stabilize, it may reflect distinct facets of skill on different attempts to perform as indicated above. Estimates of early performance are likely to measure different things on different trials for different people. Accurately separating better from poor performers or consistently determining whether a trainee has mastered a needed skill becomes difficult. This potential lack of reliability impacts the predictive power of computer-based simulations for assessing operationally oriented skills. Therefore, it affects the validity of computer simulations for job-sample performance testing in functional contexts. Widespread Navy use of computer simulations creates a need for determining the impact of the technology on the consistency and validity of computer based performance.

TECHNOLOGICAL OBJECTIVE

The technological objective of this proposed research is to conduct cognitive and statistical analyses as well as theoretical modeling to study the process of skill acquisition resulting in the stabilization of performance on a computer-based simulation of a complex cognitive task.

GENERAL APPROACH

Target Task

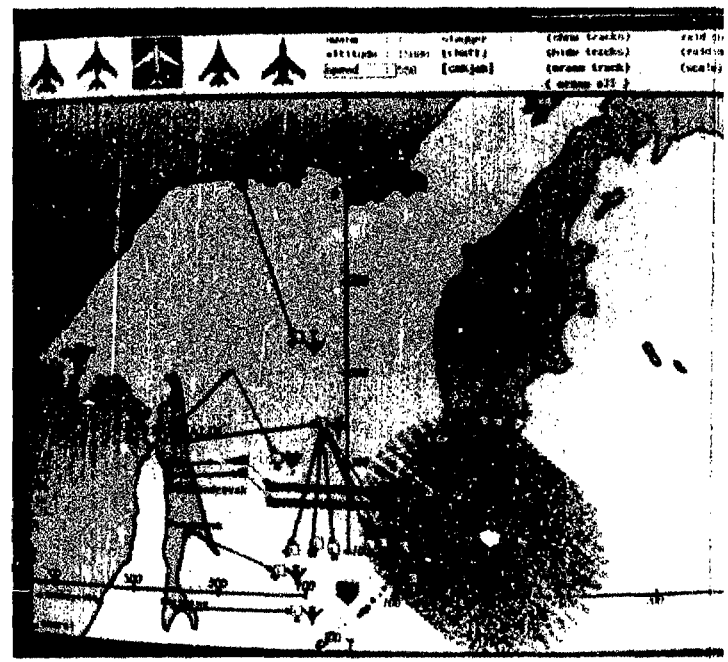
The target task of this proposed research consists of tactically allocating, deploying, and managing fighter and supporting aircraft to defend an aircraft carrier and its escort ships against threatening Soviet naval air bombers. This task demands considerable practice before it can be executed with a sufficiently high level of skill and becomes automatic. For the purposes of this research, this task is considered a test of individual differences in complex mental performance. In the execution of this task the transition from controlled to automatic performance is important. This implies that what is crucial is not early but late performance (i.e., how well individuals do after extended practice). The administration of numerous trials on this task, together with cognitive and statistical analyses, will make it possible to note when and how stabilization of performance is achieved (i.e., when the research subjects no longer show any tendency to improve or worsen with practice).

Computer-based Simulation

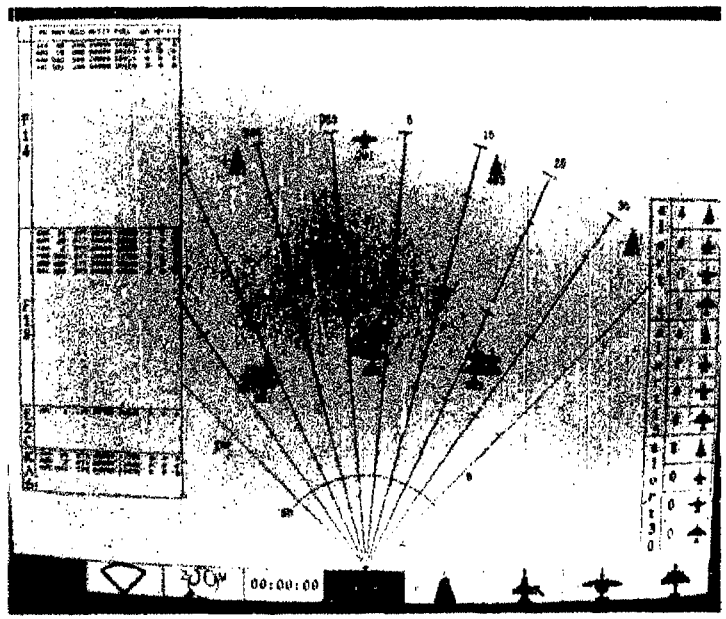
Software tools are being developed for constructing computer-based animated graphic simulations of the actual radar coverage of F-14 and F/A-18 fighters and E2-C early warning aircraft as well as fuel flow of these planes together with KA-6 tankers. These include probability of kill for Phoenix, Sparrow, and Sidewinder missiles that the different fighters carry as well. An infinite number of raids from Soviet naval air bombers with ASMs (antiship missiles) in different warfare theaters can be generated, as well as various carrier loadouts in terms of numbers of each type of fighter and missile on board. Thus, the capability exists for creating an infinite set or universe of tactical scenarios. These will be used to assess how well individuals manage outer air battles to defend carrier-based naval task forces.

Subjects

The research subjects, approximately six F-14 pilots and radar intercept officers at NAS Miramar and/or instructors and students from the Tactical Action Officer, Tactical Warfare Overview, and/or Staff Tactical Watch Officer Courses from the Fleet Combat Training Center, Pacific, will be required to allocate, deploy, and manage fighter and supporting aircraft to knock down various numbers and mixes of hostile bombers before they reach their respective ASM launch points. Each computer-based scenario will be run in compressed or accelerated time; each threat scenario will be considered as a performance test item.



Computer-generated Soviet raids using developed software.



Allocating and deploying tactical resources using developed software.

Performance Criteria

A subject's tactical performance during simulated air battles will be assessed according to 16 multivariate criteria. Some of these are as follows: The percentage of incoming threat aircraft that were detected by F-14, F/A-18, and E2-C radar systems; the percentage of bombers that fighters placed in missile launch acceptability regions (LARs); the percentage of hostile aircraft shot down or probable kills; the average range from the defended task force at which threat aircraft were knocked down; and the percentage of hostile platforms knocked down before ASMs were launched.

Procedure

Subjects will be run on the computer-based scenarios of these symbolically displayed air battles between Soviet bombers and U.S. carrier-based aircraft. How well each allocates, deploys, and manages fighters and other supporting aircraft during the simulated battle will be assessed according to the multivariate performance criteria mentioned above. The possible number of incoming raids or specific threat scenarios form a practically infinite universe. Consequently, the set of simulated tactical scenarios will be considered as an operationally oriented, domain-referenced, job-sample, performance test. With each scenario as an assessment trial, subjects will be administered 200 trials divided into 20 blocks.

Cognitive Analysis

During the first trial of every block, verbal protocols will be obtained from the subjects as they are actually conducting the simulated air battles. The analyses of these concurrent verbalizations, as well as retrospective

reports, will disclose the information heeded by the subjects while they perform this complex task. Comparisons of the thinking-aloud protocols and retrospective reports on the first trial of every block will reveal the variability in cognitive processing within as well as between subjects as they acquire skill (i.e., progress from controlled to more automatic performance of the task).

Analysis of protocols obtained early and late during practice on the task will indicate how subjects' cognitive processes and structures change as their performances tend to stabilize. These will reflect the cognitive correlates of the acquisition of stable task performance. Together with a thorough componential analysis, the information obtained from the protocol analysis will be used to construct a model for performing this complex task. This model will be used to create a theoretical framework as well as build the basis of an expert system for a computer-based "intelligent tactician" that will monitor, diagnose, and assess the conduct of simulated air battles to defend carrier task forces.

Statistical Analyses

Combining statistical procedures with protocol analyses and conceptual modeling will provide an integrated account of the cognition accompanying the acquisition of complex task performance. Together with cognitive analysis and theory, statistical techniques (e.g., a test for the homogeneity of k regression lines) can be used to uncover the mental processes and structures underlying the acquisition of stabilization.

Potential Products/Transition

The potential products of this proposed research are contributions to a knowledge base and a much needed theoretical framework. The contributions to methodology and theory culminating from this research can be naturally extended to the exploratory development of "intelligent or expert" computer-based simulation systems to measure complex cognitive performance in functional contexts. Then the predictive power of this type of performance assessment can be determined. Likewise, this follow-on work itself can be nicely transitioned to advanced development of an intelligent computer-based simulation system to support job-sample performance assessment of intricate cognitive tasks. This advanced system would allow the use of developed methodologies, theoretical orientations, mental models, as well as generic software tools to implement prescriptive procedures to aid in the production of performance tests for complex cognitive tasks.

Frederick R. Chang

Frederick R. Chang is a Research Psychologist at the Navy Personnel Research and Development Center. He received his B.A. degree in Psychology from the University of California, San Diego, and his M.A. and Ph.D. degrees in Experimental Psychology from the University of Oregon where he conducted detailed chronometric investigations of human language performance. He then took a position at Bell Laboratories in New Jersey, where he conducted research in applied cognitive psychology before taking his present position. He has published both basic and applied research in the areas of text comprehension and reading, and his current research interests are in how people interact with informational materials, as both readers and writers and ways in which technology can facilitate that interaction.

Knowledge and Process in Adult Language Competence

Frederick R. Chang **Independent Research**

The problem of adult literacy is receiving increasing attention in the Department of Defense and in the United States in general. The present research has investigated some of the factors that form the basis of reading comprehension skill in Navy enlisted personnel, with particular emphasis on understanding the difficulties among less literate individuals. The research has demonstrated the importance of prior knowledge in reading and has begun to unravel the complex interaction between "general" reading skill and content specific knowledge. Additionally, the research has developed a new approach to assessing the readability of documents, that for the first time takes into account the content knowledge possessed by the reader of the document. Aspects of this research can be applied toward the solution of persistent Navy problems connected with literacy assessment and training.

PROBLEM

The problem of adult literacy is receiving increasing attention in the Department of Defense and in the United States in general. There appear to be several reasons why the problem is surfacing now and is expected to last into the near future. The U.S. is experiencing a declining youth population. The number of 18-24 year olds available for employment will go down in the next 10 years. In 1981, for example, there were roughly 30 million 18-24 year olds available for employment; by 1995 that number will shrink to roughly 24 million. The 18-24 year olds who do not go to college form the pool from which the majority of military recruits are drawn. Today, increasing numbers of individuals in this age range are going to college. In the 1930s and 1940s, only 10-15 percent of 17-18 year olds went to college; today that figure has increased to about 56 percent.

With higher education attracting the majority of the decreasing young adult pool from the rank of the more highly educated, high-school diploma holders, the military must compete with business and industry for the talent in the remainder of the pool. To maintain current strength figures, the military will have to increase the number of high school graduates it recruits. Even though 75 percent of today's 18-19 year olds graduate from high school, a high school diploma alone does not guarantee a uniformly high level of basic skills. For example, in a representative sample of 18-23 year old youths from 1980 it was found that 75 percent had completed high school but the average reading grade level was mid-ninth grade.

The literacy problem facing the military has changed in character from some years ago. The problem today will be in training "mid-literate" personnel; people with 5th to 9th grade reading skill. The problem in World War II

concerned truly low-literate adults; individuals who lacked even the most basic reading decoding skills. The mid-literate, though, possess adequate decoding skills but lack higher level skills and the knowledge necessary to negotiate complex, technical reading tasks that are present in the military.

BACKGROUND

Adult reading education in the DoD has historically been characterized by a debate between providing general literacy training or specific literacy training. General literacy typically refers to general reading skills that can be transferred to reading materials in many different content domains. The teacher of general literacy is seen as empowering the reader with a set of reading skills that will generalize textual materials in any subject. Specific literacy, on the other hand, reflects the recognition that people read texts for specific purposes and in specific content areas. The specific literacy teacher attempts to teach the reader what he or she needs to know in order to perform a given task. Since it is generally held that it takes knowledge to get knowledge, it therefore stands to reason that if you give instruction to a reader in a content area, it will be easier for that reader to learn new information themselves in the content area.

In the Navy, the debate has been settled, at least operationally, in favor of specific literacy; it is Navy policy that all on-duty, remedial education be job-related (SECNAVINST 1510.3). Even though functional literacy is Navy policy, there are several interesting scientific questions related to the distinction between general and specific literacy. What is learned in specific literacy and how does it transfer to reading in other domains? How does the acquisition of content knowledge exert its influence in the processing of subsequent text in the same domain? In a different domain? What "generic" reading capabilities do readers possess and can these be acquired in "specific" literacy programs? Are these generic capabilities best characterized as "skills" or "knowledge?"

Recent developments in cognitive science and artificial intelligence research have pointed out the importance of knowledge in cognition. Attempts to understand the way in which humans and machines comprehend natural language have been greatly enhanced by thinking more clearly about the prior knowledge that the language comprehension device brings to the task of comprehension. Thus, modern theories of human cognition are beginning to explicate clearly the role of knowledge. These notions are slowly being incorporated into theories of reading comprehension, and are seldom used in theories of adult remedial reading.

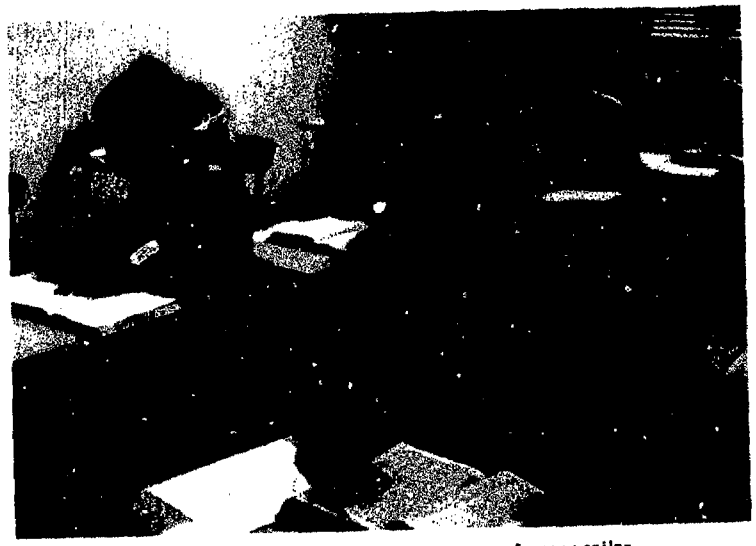
OBJECTIVE

The goal of the research is to develop a general theoretical account of the cognitive processes involved in adult literacy with an emphasis on a description of the complex interaction between prior knowledge on the one hand, and more "general" reading capabilities on the other. The research will be crucial to a clearer understanding of the difficulties facing marginally literate individuals and to techniques for improvement. At a more general level, a clearer understanding of human knowledge representation and "processing skills" that operate on that representation are at the very core of cognitive science research. The operational objectives of the research are to produce findings that will lead, in a fairly direct way, to the development of new reading assessment techniques, new "intelligent" reading tutorial techniques, and new document readability assessment techniques.

Prior knowledge of the subject area of a document enhances its readability.

FY86 ACCOMPLISHMENTS

This research project thus far has resulted in a significant scientific contribution to the field of reading and readability. Based on a developing theoretical perspective and empirical work, the research has clarified some aspects of the interaction between reading process and prior knowledge in reading comprehension in adults. The project has concentrated on the importance of prior knowledge in reading, and has conducted several experiments that show the relationship between general reading skill and prior knowledge of the content. Assuming a constant reading skill possessed by a reader, these studies have shown the "value added" to reading comprehension by having specific knowledge of the domain to be read. These findings have led to the development of a new approach to the assessment of the readability of documents that, for the first time, takes into account the prior knowledge that the reader brings to the reading task. The results show clearly that prior knowledge of the subject area greatly influences the estimated readability of a document. Additional developments include (1) a precise quantitative account of how prior knowledge (and features of the text) influence readability; (2) a determination of the influence of different reading tasks that the reader must perform on the readability of the document; and (3) a family of formulas scaled on different prior knowledge variables and different reading tasks.



Reading is an important aspect in the career of every sailor.

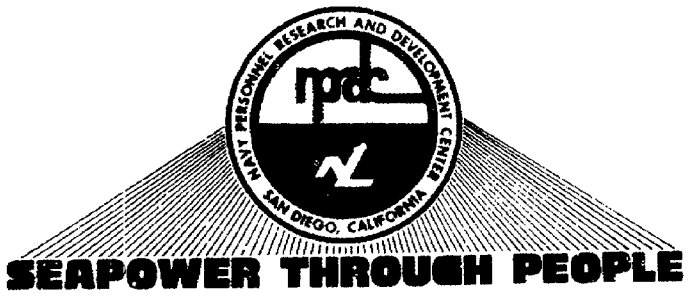
POTENTIAL PAYOFF

As mentioned earlier, currently the issue of literacy is being given high level attention in the military. The Navy in particular has addressed the issue with the recent Personal Excellence Initiative promulgated by the current and former CNO. The Personal Excellence Initiative, among other things, recommends that the Navy "significantly strengthen, expand and broaden the scope of all remedial education programs (e.g., CAMPUS, ART, JOBS) to focus on basic skills and competencies." The expansion of these remedial programs will be a major undertaking that will require a great deal of time, effort, and resources to accomplish successfully. This research will have direct payoff by providing information that will be of great value to the Initiative. Some specific benefits are described below:

a. This research has direct payoff to the area of reading assessment in the Navy. Currently, reading assessment in the Navy is conducted using standard general literacy tests. These have known problems but there is no alternative thus far. The results point up the importance of knowledge in reading asssment, and could lead to an alternative reading assessment system in the Navy. Improved reading assessment is a specific Personal Excellence Initiative.

b. This research also has direct implications for remedial reading education in showing the importance of teaching content knowledge of a technical area as separate from training "general reading skills." The Personal Excellence Initiative includes the expansion of remedial education programs in the Navy, thus the research will feed directly into decisions about how to do the expansion.

This research should have an impact on the field of readability, and consequently on how readability is assessed in the Navy. Currently, readability is assessed in the Navy using standard readability formulas. These have known weaknesses, but as yet there is no alternative. The research helps provide an improved alternative. Accurate assessment of reading materials will support the Personal Excellence Initiative.



Edwin L. Hutchins

Edwin Hutchins received his B.A. (1971), M.A. (1973), and Ph.D. (1978) degrees in Cultural Anthropology at the University of California, San Diego (UCSD). He has done anthropological field research in Papua New Guinea and aboard ships in the U.S. Navy. After receiving his doctorate degree, he took a two-year postdoctoral fellowship in Cognitive Science also at UCSD. Dr. Hutchins joined the Navy Personnel Research and Development Center (NAVPERSRANDCEN) as a Research Psychologist in 1980. His research includes the design and implementation of computer based instruction systems, and the nature of cognition in operational settings. In addition to his position at NAVPERSRANDCEN, Dr. Hutchins is also an Associate Research Cognitive Scientist in the Institute for Cognitive Science at UCSD.

Analysis of Cognition in Natural Settings

Edwin L. Hutchins

Independent Research

Most of what is now known about the human cognition is derived from experimental studies conducted in contrived settings. This project concerns the nature of human cognitive activity outside the psychologist's laboratory. A detailed study of the activities of piloting a large ship in restricted waters was conducted. The study included the collection of audio- and video-tape recordings of the activities of a navigation team composed of six individuals. A new way of thinking about the nature of human cognition and intelligence was developed and demonstrated in the analysis of the recorded data. This analysis should prove more useful than previous approaches for understanding the nature of real world task performances and predicting the consequences of changes in task settings.

BACKGROUND

Most thinking jobs in the Navy, and in nonmilitary settings as well, are accomplished by groups of people working together. In an earlier IR project, the principal investigator conducted a review of the literature in cognitive anthropology and cognitive psychology with respect to the study of naturally situated cognition. A major problem with these fields is that, as a result of assumptions deeply engrained in our own cultural tradition, they both assume as a primary unit of analysis the individual actor taken in isolation from other actors and from a world of action and mediating technologies. In the most interesting and characteristically human activities, however, the cognitive system clearly includes two or more minds, a shared culturally meaningful task, a social organization, and technologies, such as literacy, that mediate and externalize the performance of individual cognition. The cognitive properties of such systems of socially distributed cognition are certain to differ in important ways from the properties of the cognition of individuals. Yet, we know very little about how such systems actually work.

PROBLEM

An understanding of the nature of cognition in real world settings is important for diagnosing performance inadequacies, providing training that is relevant to the job to be performed, designing performance aids, and designing operational work environments. Furthermore, as work environments become more automated and include more intelligent artifacts that interact in fundamentally cognitive ways, we must develop an understanding of how systems of distributed cognition work. At present, however, a lack of appropriate theory and method hampers our understanding of cognition in settings outside the laboratory.

APPROACH

A critical step in the project was the selection of a task domain in which a system of socially distributed cognition is in operation. The system chosen was the piloting team on the bridge of a large Navy ship. This system has a number of important properties that make it amenable to study.

A small number of culturally well-defined computations are jointly performed by a group of six individuals. This means that the task is socially distributed and it is possible to know precisely what is supposed to be done, at least computationally speaking.

Each man on the bridge is a member of a decision-making team.

Several mediating technologies and representations are utilized. These include literacy, the use of charts, graphs, and other external computational media, as well as electronic devices that perform parts of the computation. The use of all of these external computational media provides additional observable evidence about the behavior of the system.

The principal investigator has more than 25 years of experience with the navigation procedures involved. This meant that the researcher could easily achieve a degree of credibility in the eyes of the team under investigation and, furthermore, that during the analysis process the researcher could concentrate on discovering the properties of the socially distributed cognitive system rather than on trying to figure out what the team was trying to do.

The most fundamental theoretical innovation of the approach taken here is a change in the boundaries of the cognitive system that is to be the unit of analysis. Rather than looking primarily at the individual actor and then trying to see how a set of individuals could jointly function, this project takes a group of people jointly performing a single task as the fundamental unit. Communication among the actors is seen as an internal process of the cognitive system. Computational media, such as charts, are seen as representations internal to the system, and the computations carried out upon them are more processes internal to the system. Yet, because the cognitive activity is distributed across a social network, these internal processes and internal communications are directly observable. If a cognitive psychologist could get inside a human mind, he or she would want to look at the nature of the representation of knowledge, the nature and kind of communication among processes, and the

organization of the information processing apparatus. We might imagine in such a fantasy that, at some level of detail, underlying processes would be obscured. If these observations could be made, however, any cognitive psychologist would be very happy. With systems of socially distributed cognition, we CAN step inside the cognitive system. The analysis in this project begins by attending to precisely those observations the cognitive psychologist would attend to were he or she given the opportunity to step inside a mind.

The data consist of written, audio- and video-taped observations of the behavior of the piloting team on the bridge of a Navy ship in high workload situations. A true crisis situation and the system's response to it were also observed and recorded. Data were collected during two at-sea periods in February and April 1985. The transcribed data have been analyzed using both qualitative and quantitative methods to answer the following questions: How is computation distributed across space, time, people, and artificats in this task? How is coordination among the activities of several actors maintained? How do the computational properties of such systems differ from those of the individuals who participate in them? What is the nature of on-the-job learning? How does the system adapt to changes in the operational environment?

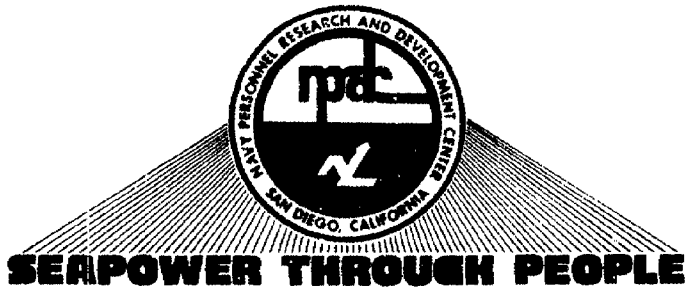
RESULTS

This project has developed a new theory of naturally situated cognition and has applied this theory to a particular instance of naturally situated cognition: large ship piloting. Some of the results are specific to the domain of navigation. The description of the activities of piloting a large ship in restricted waters, for example,

On the bridge, one can see the "mind" of the ship in operation.

form a specific case that should be of interest to other researchers interested in performance in that domain. The transcripts of the data themselves are voluminous and represent approximately one person-year of effort. They are a unique corpus of task embedded discourse.

The intent of the project has never been simply to describe a particular task. Instead, the project has tried to work through a different way of thinking about the nature of human cognition and intelligence in general; one that is applicable in behavioral contexts that are organized by a culture rather than by a psychological experimenter. This project has developed a different way to anticipate the consequences of changes in manning or of the introduction of new technology to navigation and other socially distributed task settings. The new theory also interacts with other ongoing NPRDC projects. In one case, it has already spawned a new human/machine interface design.



Leonard J. Trejo and test subject.

Leonard J. Trejo was born in Mexico City, Mexico on 24 February 1955. He did his undergraduate studies at Lock Haven State College in Pennsylvania and later at the University of Oregon, receiving the B.S. (1977) degree in psychology. He did graduate work in psychobiology at the University of Michigan and later at the University of California, San Diego, with emphasis on sensation and perception. He received M.A. (1980) and Ph.D. (1982) degrees in psychology from the University of California for research on the neurophysiological control of the pupillary light reflex and on visual sensitivity loss in hereditary retinal degeneration. From 1982 to 1984, he served as a Senior Fellow in the Department of Ophthalmology at the University of Washington where he performed research on retinal toxicity, physiology of color vision, and neuroanatomy of the visual cortex. Since September 1984, he has been a Personnel Research Psychologist at the Navy Personnel Research and Development Center. His research interests include human color vision, evoked potentials and magnetic fields, and biopsychometric assessment.

Brain Mechanisms for Human Color Vision

Leonard J. Trejo and Gregory W. Lewis

Independent Research

We are studying neuroelectric (evoked potential) measures of human color vision mechanisms to test their applicability to personnel assessment and human factors engineering. Our approach is to use the technique of color exchange stimulation, which involves changing only the color of a visual stimulus over time, while holding the size, shape, position, and texture constant. Using a computerized video stimulator, we recorded evoked potential (EP) responses to eight different color exchange stimuli in six subjects with normal color vision. These EPs suggest that there are individual and day-to-day variations in brain processing for color. We discuss how knowledge of these variations may improve personnel assessment and human factors engineering tasks, such as the design of color-coded display systems.

BACKGROUND

Much experimental evidence from the last three decades has led to the view that color vision is three-dimensional and that its three dimensions are subserved by three distinct brain mechanisms. These include two chromatic mechanisms, red-green (R-G) and blue-yellow (B-Y), and one achromatic (A) or black-white mechanism.

The focus in the current research is on the chromatic mechanisms for two reasons. First, their activity is thought to determine the perceived hues in a visual scene. For example, nobody ever perceives a color as being both reddish and greenish at the same time. Similarly, nobody ever perceives a color as being both bluish and yellowish at the same time. The mutual exclusiveness of color perceptions is thought to arise from the opponent nature of the chromatic mechanisms. In the case of the R-G mechanism, opponent process theory proposes that stimulation of the mechanism in one direction (e.g., neural excitation) produces a sensation of redness, whereas stimulation in the other direction (e.g., neural inhibition) produces a sensation of greenness. At some balance point between redness and greenness the mechanism is nulled or inactive and produces no sensation at all. When the R-G mechanism is nulled, only the B-Y mechanism produces a hue sensation. This is why a balanced mixture of red and green light produces yellow. The mixture of red and green light falls into a region of the spectrum that the B-Y mechanism responds to by producing a sensation of yellowness. The R-G mechanism does not contribute to the sensation of yellowness. The second reason for studying the chromatic mechanisms is that they mediate chromatic discrimination, which is the ability of the visual system to detect and respond to hue differences independently of intensity differences. Chromatic discrimination helps us distinguish colored symbols on colored backgrounds, such as those that appear on many computer displays.

Hue perception and chromatic discrimination vary among individuals and across stimulus conditions. Variations also exist within an individual on a day-to-day basis and may reflect stress, fatigue, drug, or other biochemical effects. Such variability may provide new information for personnel assessment, display systems engineering, and for understanding the basic physiology of color vision.

Research suggests that the mechanisms proposed by the opponent process theory are useful for explaining variability in color vision. Recent work by Boynton and Kambe (1980) and Nagy, Eskew, and Boynton (1986) shows that individual differences in human color discrimination can be explained by equations that represent the R-G, B-Y, and A mechanisms. The equations account for individual differences by adjustment of a set of constants. Thus, the form of each equation captures the essence of a mechanism, and the constants fit the equation to the performance of an individual. The generality of the equations suggests that the basic organization of the visual system is qualitatively similar across individuals, but the variability of the constants suggests that certain factors vary quantitatively from one individual to another.

OBJECTIVE

The goal of this research is to identify physiological measures of human color vision mechanisms proposed by the opponent process theory and described by the equations developed by Boynton and co-workers. Procedures include recording visual evoked potentials

Individual variability in brain processing of color is important for the design of color-coded display systems.

(EPs) obtained from stimulating the mechanisms with computerized visual displays.[1] Others have shown that EPs are strongly influenced by the activity of color mechanisms (Kinney, McKay, Mensch, and Luria, 1972; Krauskopf, 1973; Regan, 1981; Riggs & Sternheim, 1969; White, White, & Hintze, 1986). We will attempt to identify properties of EPs that provide new information, which can be used to improve personnel assessment and human factors engineering. Properties that show large individual differences may relate to structural and functional differences in the visual systems of different subjects. Such differences may influence the processing of color-coded information. We will also attempt to relate properties of color-evoked EPs to the basic physiology of human color vision. Properties that are relatively stable over time and across individuals may allow inferences about the basic structure and function of the human brain.

APPROACH

Exchange Stimulation

The method of exchange stimulation is used for selectively activating mechanisms of color vision (Estevez & Spekreijse, 1982). The method involves changing only the color of a visual stimulus over time, while holding the size, shape, position, and texture constant. By holding intensity (luminance) constant, it is possible to activate a single chromatic mechanism or a particular combination of the two chromatic mechanisms without affecting the achromatic mechanism. For example, the intensities of a red light and a green light can be adjusted so that the exchange of one light for the other will stimulate only the R-G mechanism. In this case the lights have equal stimulating effects on the B-Y and A mechanisms but unequal stimulating effects on the R-G mechanism. Any neural response that occurs as a result of this exchange must be due to the R-G mechanism acting alone. Similarly, the exchange of properly adjusted blue and yellow lights will stimulate only the B-Y mechanism. Other pairs of lights, orange and magenta (bluish red) for example, stimulate both the R-G and B-Y mechanisms about equally and can also be nulled for the A mechanism.

We have developed equations for transforming the intensities of the three video monitor primary colors (red, green, blue) into excitation values for the R-G, B-Y, and A mechanisms. These equations are based on visual sensitivity data (Smith & Pokorny, 1975) and a colorimetry system developed specifically for work with chromatic mechanisms (MacLeod & Boynton, 1979). The video monitor was calibrated using methods suggested by Cowan (1983). We also developed a software system for performing the calculations and transformations necessary to produce color exchange stimuli. Details of our stimulation system will appear in a forthcoming report (Trejo, Nagy, & Lewis, in preparation).

Initially, we selected eight pairs of exchange stimuli that relate to prior psychophysical data on chromatic discrimination (Boynton & Kambe, 1980). These stimuli

[1] EPs are very small voltage signals (microvolts) recorded from electrodes placed on the scalp that represent the response of the brain to sensory input. EPs must be extracted from larger ongoing electroencephalographic (EEG) activity by signal averaging.

were designed so that, for a normal subject, the perceived difference between the members of one pair would be roughly equal for all eight pairs. For example, the perceived difference between the red and green members of one pair would be roughly the same as the perceived difference between the blue and yellow members of another pair. The stimuli were circular and centrally fixated, subtending three degrees of visual angle with an exchange rate of 1 hertz (square wave). The stimulus surround was neutral white (D6500, a chromaticity reference point) and rectangular (15 degrees horizontally, 10 degrees vertically). Each member of the eight pairs of exchange stimuli and the surround were set to a luminance of 4.41 foot-lamberts.

Electrophysiological Recording

EPs were recorded using commercially available nylon helmets (Electrocap International[2]) fitted with tin electrodes. Recording sites were F3, T3, P3, O1, F4, T4, P4, and O2 (Jasper, 1958), referenced to the nose. Signals were amplified (20,000 times) and band-pass filtered between 0.3 and 30 hertz (Grass Model 12 Neurodata Acquisition System). The sampling rate was 120 hertz. EP records were monitored for eye blink or eye movement artifact. Fifteen cycles of EP data were used to produce EP waveform averages for each of the eight stimulus conditions. Our dependent measure was standard deviation amplitude, which is an integrated measure of the total amplitude in a waveform and is proportional to the square root of the power.

[2]Identification is for documentation only and does not imply endorsement.

A red versus green exchange was presented to the subject and was followed by magenta versus greenish-yellow, blue versus yellow, cyan versus orange, and four pairs of intermediate hues. The red versus green and blue versus yellow exchange stimuli were designed to only activate the R-G and B-Y mechanisms respectively. All other exchange stimuli activated combinations of the mechanisms. None of the stimuli activated the A mechanism.

RESULTS

Averaged EPs from an electrode over the left occipital region (O1) of subject GWL for each of the eight pairs of test stimuli are shown in Figure 1. Each of the sixteen EPs represents the average brain activity for half a cycle of an exchange. The waveform along the axis labeled "red" represents the average EP evoked by 15 stimulus transitions from green to red. Across the diagram, the waveform along the axis labeled "green" represents the average EP for 15 transitions from red to green, or the average activity during the green phase of the stimulus.

The plot in the center of Figure 1 describes the variation in the EP amplitude of the average EPs as a function of the relative excitation of the chromatic mechanisms. The distance from the center (0, 0) to each point in the plot is the standard deviation amplitude of the average EP in a specific direction. The lines connecting the points in the plot form a chromatic

response contour that relates the amplitude of the EP to stimulation of the chromatic mechanisms. For example, the two contour points lying on the x-axis reflect the EP amplitudes produced by stimulating the R-G mechanism either towards increasing redness (right) or increasing greenness (left). Similarly, the two points on the ordinate reflect stimulation of the B-Y mechanism in opposite directions. Points along the diagonals represent combined activation of the two mechanisms. Subject GWL showed generally greater reactivity to exchanges along the blue-yellow dimension than along the red-green dimension.

Chromatic response contours for subject GWL were measured on two occasions separated by a week and are compared in Figure 2. Figure 2 also plots the chromatic response contours for five other subjects that were measured on single occasions. The two contours of subject GWL are similar, particularly in the lower two quadrants, which represent greens, yellows, oranges, and reds. However, there are differences between the contours in the upper quadrant, where sensitivity to blue seems to be depressed on the second run as compared with the first run.

Each of the six subjects in Figure 2 has a unique chromatic response contour. In terms of the primary mechanisms, subjects GWL, AJC, and DC show relatively greater blue-yellow than red-green responses, whereas subject ELF clearly has greater red-green than blue-yellow responses. GC shows about equal responses to both mechanisms. The EPs of the six subjects also differ in their overall sensitivity to chromatic exchanges as shown by the total area enclosed by the chromatic response contours. In particular, the EPs of subject GWL are generally larger than those of the other five subjects.

CONCLUSION

This research demonstrates that both individual and day-to-day variations in sensitivity to color exchange stimulation occur. Clearly, we must measure these variations more closely and test their significance. Our experience suggests that the differences seen in Figure 2 are reliable. Therefore, we conclude that EPs evoked by exchange stimulation may provide unique data that describe individual and subtle day-to-day differences in brain processing of color.

Individual variability in brain processing of color is undoubtedly important for the design of color-coded display systems. With data such as those provided by Figure 2 one can begin to examine how, for example, a display system could be customized for optimum brain processing of color codes by different individuals. Our experience in other areas has also shown that EP variability correlates with factors such as on-job performance and personnel reliability (Lewis, 1983; Lewis, Trejo, Blackburn & Blankenship, 1986).

We are now examining the relationships between color EP variability and job performance. Exchange stimulus technology has become part of our standard EP test battery. Color EP data are being acquired for two related research projects, Future Technologies for Manpower and Personnel Systems, and Marine Corps Biopsychometrics. Both use USMC NCO students at the Division Schools, Camp Pendleton.

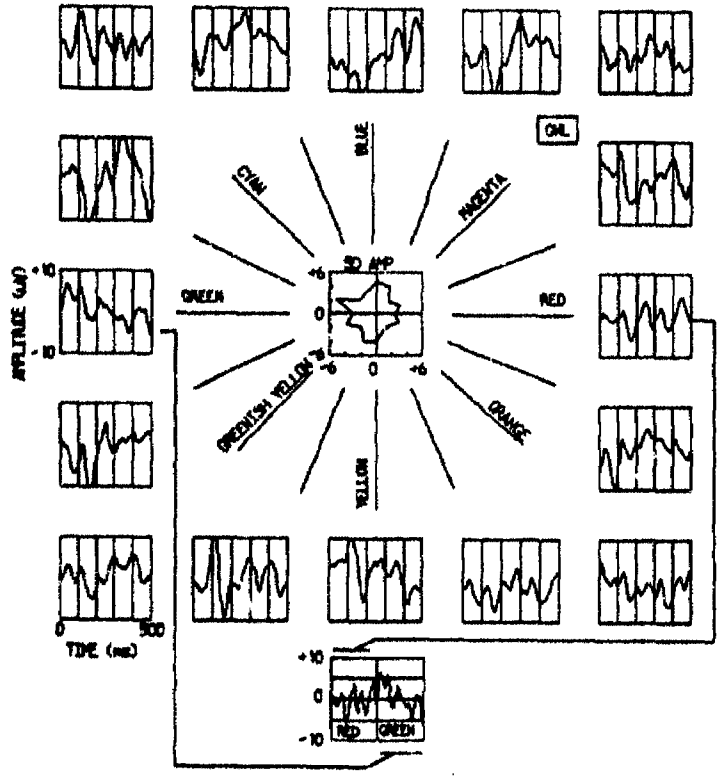


Figure 1. Shown are averaged EPs from the left occipital electrode (01) for subject GWL in response to eight different color exchange stimuli. Each of the sixteen waveforms is the EP to one half cycle of the exchange along the indicated axis. The period of each half cycle is 500 ms; vertical bars indicate 100-ms time increments. Amplitude scale is in units of 10^{-6} volts (μ V). The inset below illustrates how the two half cycles along the red-green axis are taken from one complete cycle of the exchange stimulus. The center graph shows the standard deviation amplitude in μ V X 10 for the sixteen half cycle EPs as a function of stimulus direction. Horizontal and vertical axes correspond to stimulation of isolated R-G and B-Y mechanisms respectively. Oblique axes correspond to combined stimulation of the two mechanisms. Color names of major stimulus directions are labeled.

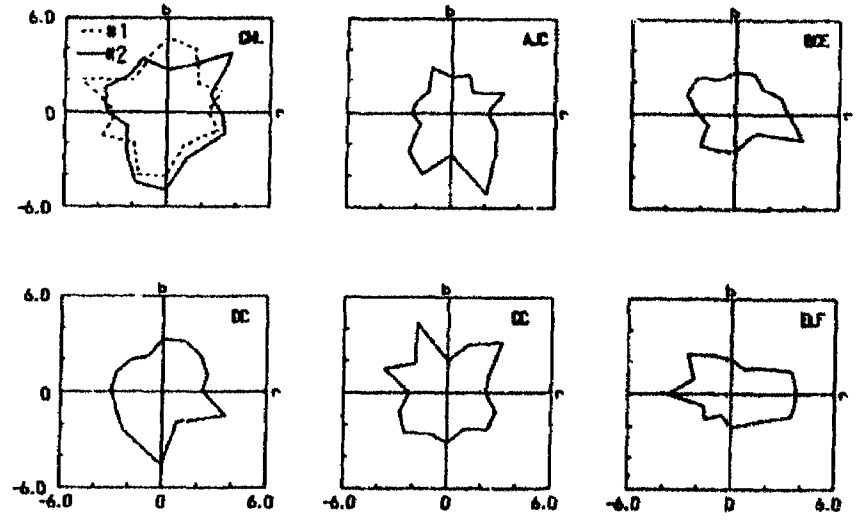


Figure 2. Chromatic EP response contours for subject GWL measured on two occasions separated by a week and for five other subjects measured on single occasions. The conventions of the plot are described in Figure 1. All six subjects have normal color vision.

RECOMMENDATIONS

Future research will seek quantitative descriptions of individual differences in color processing and the relationship of these differences to personnel assessment and human factors engineering. Rigorous definitions of sensitivity to color exchange will be sought.

Another interest is in localization of neuronal sources subserving color mechanisms. For this work we plan to record evoked magnetic fields. Such recordings measure a different dimension of brain activity, the magnetic field generated by neuronal electric activity; they may also provide improved source localization. In related research efforts, we are developing new methods for removing reference electrode and volume conduction effects from neuroelectric recordings (Nunez, 1986a, 1986b). Such procedures will be integrated into our color research.

REFERENCES

Boynton, R. M., & Kambe, N. (1980). Chromatic difference steps of moderate size measured along theoretically critical axes. Color Research and Application, 5, 13-23.

Cowan, M. (1983). An inexpensive scheme for calibration of a colour monitor in terms of CIE standard coordinates. Computer Graphics, 17, 315-321.

Estevez, O., & Spekreijse, H. (1982). The "silent substitution" method in visual research. Vision Research, 22 681-691.

Jasper, H. (1958). The ten-twenty electrode system of the International Federation. Electroencephalography and Clinical Neurophysiology, 10, 371-375.

Kinney, J. S., McKay, C. L., Mensch, A. J., & Luria, S. M. (1972). Techniques for analysing differences in VERs: Colored and patterned stimuli. Vision Research, 12, 1733-1747.

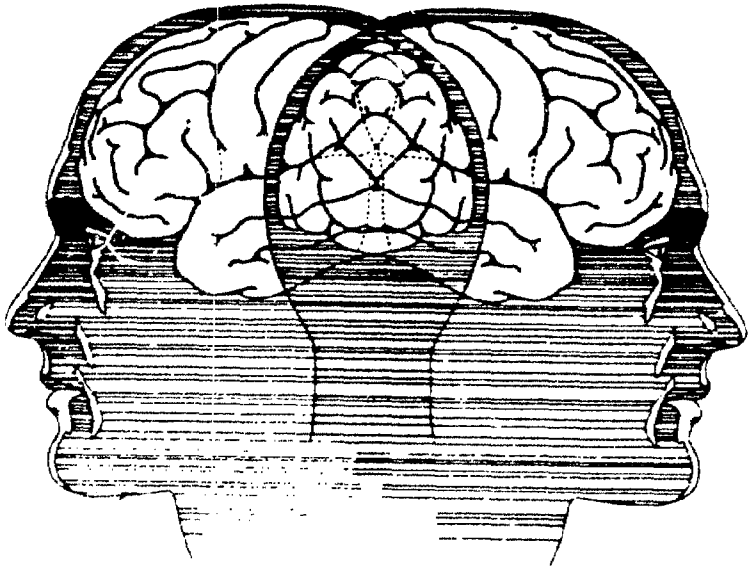
Krauskopf, J. (1973). Contributions of the primary chromatic mechanisms to the generation of visual evoked potentials. Vision Research, 13, 2289-2298.

Lewis, G. (1983). Event related brain electrical and magnetic activity: Toward predicting on-job performance. International Journal of Neuroscience, 18, 159-182.

Lewis, G. W., Trejo, L. J., Blackburn, M. R., & Blankenship, M. H. (1986). Neuroelectric and neuromagnetic recordings: Possible new predictors of on-job performance. Proceedings in the Department of Defense, Tenth Symposium. USAFA-TR-86-1, 606-610.

MacLeod, D. I. A., & Boynton, R. M. (1979). Chromaticity diagram showing cone excitation by stimuli of equal luminance. Journal of the Optical Society of America, 69, 1183-1186.

Nagy, A. L., Eskew, R. T., & Boynton, R. M. (1986). Color discrimination contours in a cone excitation diagram. Unpublished manuscript.



Nunez, P. L. (1986a). Removal of reference electrode and volume conduction effects from evoked potentials: I. Derivation of method and computer simulations (HFOSL Tech. Note 71-86-05). San Diego: Navy Personnel Research and Development Center.

Nunez, P. L. (1986b). Locating sources of the brain's electric and magnetic fields: Some effects of inhomogeneity and multiple sources, with implications for the future (HFOSL Tech Note 71-86-12). San Diego: Navy Personnel Research and Development Center.

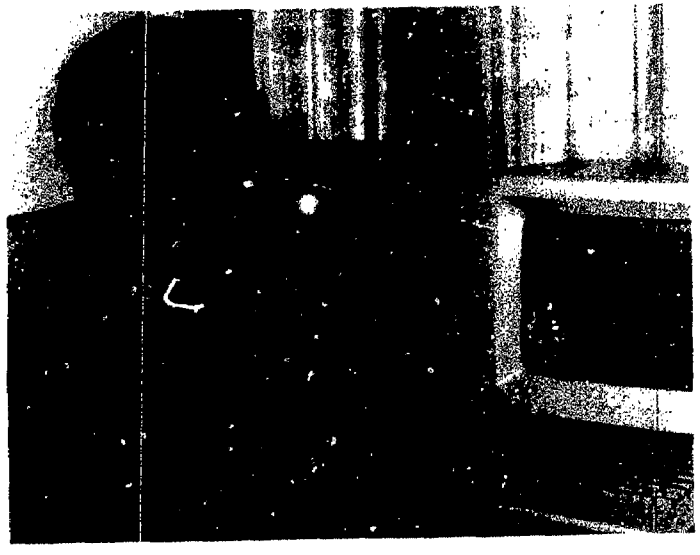
Regan, D. (1981). Evoked potential studies of visual perception. Canadian Journal of Psychology, 35, 77-112.

Riggs, L. A., & Sternheim, C. E. (1969). Human retinal and occipital potentials evoked by changes in the wavelength of the stimulating light. Journal of the Optical Society of America, 59, 635-640.

Smith, V. C., & Pokorny, J. (1975). Spectral sensitivity of the foveal cone photopigments between 400 and 500 nm. Vision Research, 15, 161-171.

Trejo, L. J., Nagy, A. L., & Lewis, G. W. (in preparation). Brain mechanisms of human color vision: I. Evoked potentials produced by chromatic exchange stimulation (NPRDC Tech. Rep.). San Diego: Navy Personnel Research and Development Center.

White, C. T., White, C. L., & Hintze, R. W. (1986). Temporal interactions among the color and pattern-specific components of the visual evoked potentials (HFOSL Tech. Letter 71-86-04). San Diego: Navy Personnel Research and Development Center.



Louis M. Weitzman

Louis Weitzman received a Bachelor of Architecture from the University of Minnesota in 1974 and a Master of Architecture in Advanced Studies from Massachusetts Institute of Technology in 1978. At MIT he worked with the Architecture Machine Group concentrating in computer-aided design with an emphasis in color raster graphics systems. After graduating, Louis helped to design and implment a number of computer systems including an energy accounting package analyzing the energy performance profiles of commercial buildings, and a highly interactive sketch/paint system in conjunction with an architectural computer-aided design system. Since coming to the Center, he has helped to build training systems using techniques from the area of artificial intelligence. His interests include building intelligent tools to aid in the process of design.

Designer: A Knowledge-Based Graphic Design Assistant [1]

Louis M. Weitzman

Independent Exploratory Development

Designer is an interactive tool for assisting with the design of two-dimensional graphic interfaces for computer-based instructional systems. The underlying motivation is to improve the quality of the interfaces by making them more consistent and visually more effective. Graphic domain knowledge, stored in a frame-based representational facility, is coupled to a domain independent mechanism to analyze and critique the user's original design, and then synthesize design alternatives. These alternative solutions are generated within a design context, or style, and are based upon graphic constraints.

OVERVIEW

Applying technologies from artificial intelligence and cognitive science to the development of computer-based training and computer-aided design systems can provide support in areas where developers and users lack expertise. In addition, intelligent tools can substantially enhance the process of design. Designer is an interactive tool for the users of Steamer's Graphics Editor. Steamer is a computer-based training system that is used to help students develop an understanding of the complex domain of steam propulsion (Hollan, Hutchins, & Weitzman, 1984). Steamer consists of a color graphics interface to a mathematical simulation and a black-and-white control screen. Users can view and manipulate this simulation at a number of different hierarchical levels through the color interface. The current system contains over 100 color views that range from abstract, high-level representations of the propulsion plant to detailed views of gauge panels quite like the actual gauge panels in a ship. The Graphics Editor allows nonprogrammers to create interactive, dynamic views of the simulation graphically. This facility has allowed propulsion engineering instructors to create substantial portions of the student interface to this advanced training system. Even though the Graphics Editor was developed to construct Steamer views, it is domain independent and has been used in a variety of domains including monitoring the real-time performance of a computer operating system and the control of a video production switcher.

[1] Previously titled "Development of Graphic Design Aids."

The user constructs the views from graphic components called icons, which represent elements in the application domain. Icons perform two tasks. First, they graphically depict the state of the simulation. Second, they enable the user to affect the simulation by positioning a cursor over the icon and clicking the mouse. In creating a view, the user selects the icons to be added to the view from a menu on the black-and-white screen. The user then positions and sizes the new icon on the color display. The icon's parameters default according to the type of icon chosen. Then, through a process of incremental refinement, the user modifies only those attributes that differ in this particular application.

Facilities were built into the Graphics Editor to support the construction of graphically consistent views. These facilities include the constraints of various types of grid latching and the maintenance of aspect ratio for specific icons. However, these constraints were often overridden by the designer. Even working within these constraints, users of the Graphics Editor often violate important graphic design principles and have difficulty maintaining stylistic conventions across sets of views. It is unrealistic to assume that instructional designers are knowledgeable about graphic design. Designer is a tool to enhance the Graphics Editor by supplementing the designer's domain knowledge with the necessary graphic expertise.

APPROACH

Designer provides visual expertise to users of the Graphics Editor interactively constructing new views or modifying existing ones. It consists of three interrelated processes, an Analyzer, a Critiquer, and a Synthesizer, coupled to a domain-dependent knowledge base. This knowledge base consists of design elements, design relations, techniques for their identification, sets of constraints for establishing a context--or style--for critiquing a design, and generative techniques for creating design alternatives. The Analyzer first parses a design based on the elements and relations of the domain and records this information in the knowledge base. The Critiquer uses this information with domain-based design constraints to indicate where the current design succeeds or fails. Using knowledge representing the current state of the design, the Synthesizer then generates design alternatives within a design style. The separation of these three processes from the knowledge base provides independence and modularity to the system and creates a technology that can be extended to other design domains.

To support the internal mechanisms of Designer, two generic software tools have been incorporated into the system architecture. These tools are Steamer's frame-based knowledge representation facility (MSG) for storing all domain knowledge and an assumption-based truth maintenance system (ATMS) for maintaining alternative design decisions that define the design space.

DOMAIN-DEPENDENT KNOWLEDGE REPRESENTATION

Much has been written about the knowledge required for graphic design. Unfortunately, the literature does not suggest a consistent representation for this knowledge. Designer incorporates much of this knowledge and maintains it in the frame-based representational system, or MSG. Designer concentrates on the knowledge describing the domain elements, their relations, graphic constraints imposed on both the elements and their relations, and graphic techniques for their modification within an existing style.

Elements

For general graphic design, the domain elements refer to points, lines, planes, etc. In Designer, however, these domain elements represent the icons contained within a Steamer view which are characterized by their graphic properties. The MSG class of elements records all domain elements that will be used in subsequent design analyses. The attributes of this MSG class include graphic properties--color, size, location, type, and shape--that distinguish these elements. The values of these properties on an instance are in fact instances of other MSG classes that represent valid values for the class. For example, the class of color includes instances for Steamer's basic colors. The class of size includes instances describing a range of sizes from very small to very large, while the class of shape includes instances of basic geometric shapes such as linear, circular, rectangular, etc.

Relations

Currently, the graphic relations in the knowledge base are similarity, proximity, grouping, and repetition. The relations of similarity, grouping, and repetition are further classified by the graphic properties of the elements (e.g., grouping by color, repetition by type, etc.). All relations know how to handle a generic message in order to identify occurrences in the design of the relation that they represent. When an occurrence is identified, a new instance of the class is created, stored on the class object, and initialized with all the elements participating in the relation. Relations can also build on one another. For example, elements in proximity to one another may form grouping relations, and groupings may form repetition relations (depending on the elements properties and their layout).

Constraints

Domain constraints consist of both basic graphic design principles, which are important in the construction of two-dimensional views, and graphic design standards, which are adopted for the current application. Principles are constraints that are true for all graphic designs and are generally accepted methods of making views consistent, unambiguous, and visually effective. The principle of significant difference of size, for instance, states that when elements differ in size, their size should be significantly different to avoid a sense of ambiguity (Sherwood, 1981). Graphic design standards are special constraints that exist for a given set of designs for a given application. In Steamer, the use of a title is a standard. The restriction that pipes be within an acceptable range of sizes is another standard.

Styles

A design should be sensitive to the context in which it is created. This context defines the external constraints that shape and guide the final solution. In Designer, this context is referred to as a style and is constructed by selecting the constraints (principles and standards) that are to be enforced within this context. Good design in one style may not necessarily be good design in another. Modifying the style within which a critique is made ultimately affects the final form of the design. A graphic style is also defined by the visual techniques employed in the communication of information These visual techniques represent a vocabulary in which to describe the design. The techniques used in conjunction with the constraints may suggest a variety of graphic procedures to modify an alternative. The constraints indicate a discrepancy in the design, while the interaction of the techniques suggest the graphic procedures--possibly more than one--that will modify the design.

DESIGNER'S PROCESSES

Design involves a cycle of gathering information, critiquing that information, and making decisions based on those critiques. New information gleaned from this process is incorporated back into the cycle for subsequent refinement of the design. This analysis-critique-synthesis cycle is a general process used in all design whether for computer interfaces, industrial applications, or architecture. The process is domain independent. Designer carries out the analysis phase of the design process by parsing the design into domain elements and relations. Second, the system must locate areas that need to be improved; this is Designer's critique process. Once these two steps have occurred, the system is ready to suggest alternatives for modifying the design; this is Designer's synthesis process.

Analysis

The analysis process parses the design and locates existing domain elements and relations. Identifying the elements is straightforward because of the use of icons in the Steamer interface and the object-oriented nature of their implementation. An instance of the MSG class elements is created for each icon. The attributes of this element are appropriately initialized with values for each property being an instance of the corresponding MSG class. Once the domain elements have been created, the system locates instances of domain relations. To maintain the independent nature of the analysis, generic messages are sent to each relation class to identify instances of the class within the design. When an occurrence is found, an instance of the MSG relation class is created and initialized. This includes the recording of the elements that participate in the relation on the appropriate attribute of the MSG object.

Critique

As Christopher Alexander suggests, the notion of a misfit is more compelling than that of a fit, and it is a driving force behind the ultimate shape of a design (1974). In Designer, the misfits are identified as violations of the domain constraints. The Critiquer creates a comment for each unsatisfied design constraint within the current style. Descriptions and justifications of the comments that are based on this constraint can then be presented. These comments are displayed in the scrolling pane of the black-and-white interface, are mouse sensitive, can be highlighted, and can be described in terms of their underlying principle or standard. Under this paradigm, it becomes possible to request multiple critiques, each based on a different style. This is an especially powerful paradigm for views that may need to be presented in different media, each with different constraints. For example, a style appropriate for a high-resolution color display may not be appropriate for a black-and-white hardcopy presentation where features are not as clearly distinguishable.

Synthesis

Design decisions are made in the synthesis phase to refine the elements and their relations incrementally. Knowledge of the elements, their relations, and the comments from the critiquing phase forms the basis for these design modifications. Each comment communicates via generic messages to the constraint on which it is based in order to determine the procedures for satisfying the constraint. More than one procedure may be available to satisfy the constraint and all possibilities are presented to the user. These procedures are a result of the interaction of the various visual techniques and the design constraints that describe the style.

The design alternatives are maintained by a new form of truth maintenance system, or ATMS (De Kleer, 1986a, 1986b, 1986c). This new form of truth maintenance system is well suited for simultaneously tracking multiple alternatives in the design space where a reasonable number of the potential solutions must be examined. Unlike previous truth maintenance systems that only manipulated assumptions, this system also manipulates sets of assumptions. As a result, inconsistent information can exist and it is possible to work effectively and efficiently in the problem space. Context switching is free, and most backtracking and all retraction is avoided.

In Designer, the assumptions that are manipulated are the alternatives created by incremental design decisions. Solutions at any stage in the design process are the consistent, noncontradictory environments maintained by the ATMS. Any contradictions that arise are handled by the ATMS and will not appear in the same environment. The user can explore the design space by interactively inspecting the individual ATMS environments. Each solution can be displayed on the color screen and can explain itself in terms of the underlying assumptions and justifications. Based on these assumptions and justifications, an alternative can describe its derivation and individual decisions can be described in terms of their potential contribution to a final solution. The system then conveys design precepts while the user is viewing a specific design alternative. Thus, the user's knowledge of constructing visual presentations is enhanced for future design.

CONCLUSION

The initial implementation of Designer has been completed (Figure 1) and used on existing Steamer diagrams. It has provided useful feedback and encouraging results. Even in views that were considered to have been carefully crafted, the system noted inconsistencies and suggested improvements. In addition, a preliminary use of the ATMS has shown the feasibility of using it to maintain solutions in the design space.

As Designer finds its way into the daily use of the Graphics Editor, the quality of the interfaces constructed and the subsequent instruction in which they are used will improve. Additionally, the generic approach taken results in the understanding and construction of a technology that can be utilized throughout the laboratory. These tools not only include the knowledge representation facility, MSG, and the ATMS, but also include the use of the generic design strategy available to other domains.

PROJECT TRANSITION

The feasibility of the approach and subsystems utilized in Designer have been demonstrated and suggests the transition of Designer into other work. Designer meshes nicely with other projects because of the common goal of building powerful tools to help create sophisticated environments for future training needs. Designer will be transitioned into the 6.3 project, "AI Tools for Development of Instruction," in FY87.

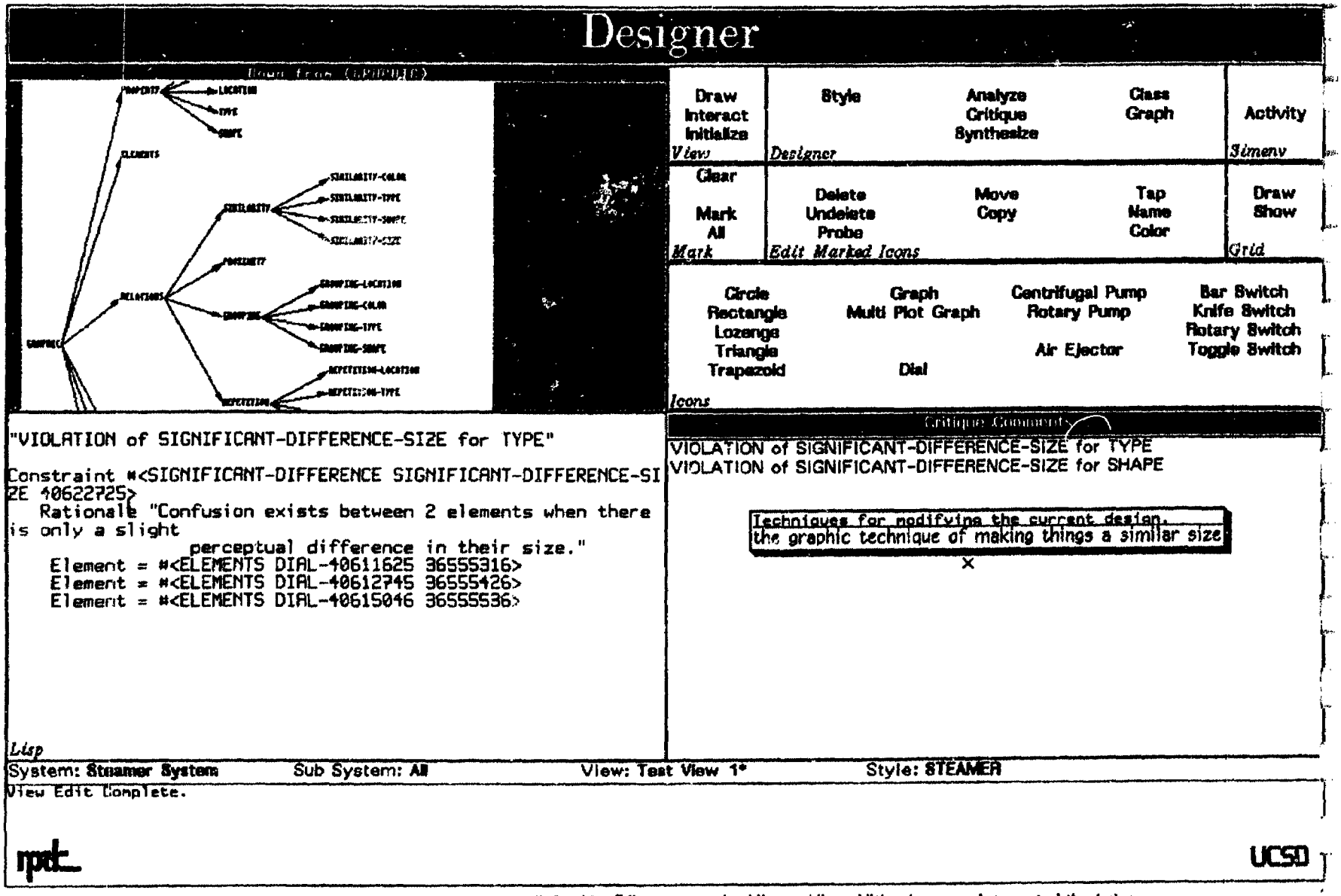


Figure 1. The Designer interface provides access to all Graphics Editor commands while providing additional commands to control the design processes and related functions. Domain-dependent knowledge is displayed in a graphing window pane. State information (e.g., current values for system, subsystem, view, and style) is provided in the status line near the bottom of the screen.

REFERENCES

Alexander, C. (1974). Notes on the synthesis of form. Cambridge, MA: Harvard University Press.

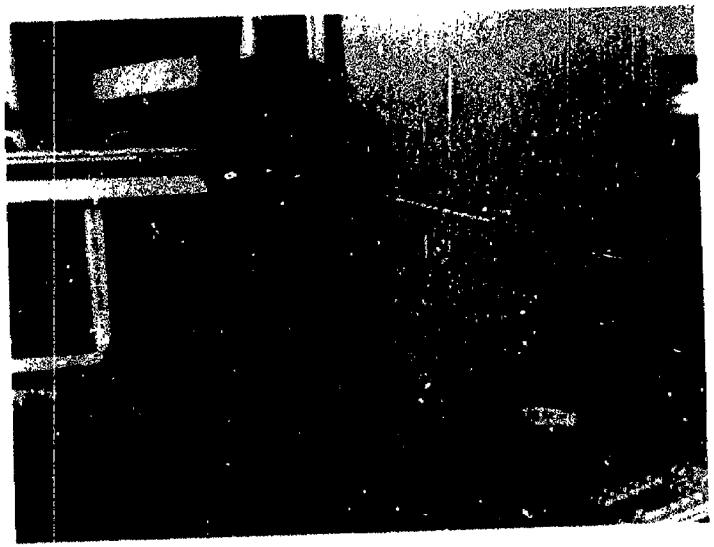
De Kleer, J. (1986a). An assumption-based TMS. Artificial Intelligence, 28, 127-162.

De Kleer, J. (1986b). Extending the ATMS. Artificial Intelligence, 28, 163-196.

De Kleer, J. (1986c). Problem solving with the ATMS. Artificial Intelligence, 28, 197-224.

Hollan, J., Hutchins, E., Weitzman, L. (1984). Steamer: An interactive inspectable simulation-based training system. AI Magazine, 5(2), 15-28. Also to appear in Greg Kearsley (Ed.), Artificial intelligence and instruction: Application and methods. Reding, MA: Addison-Wesley.

Sherwood, R. (1981). Principles and elements of architecture. Los Angeles, CA: University of Southern California Press.



Donald B. Malkoff

Dr. Donald B. Malkoff majored in physics as an undergraduate at Harvard University. He received an M.D. degree from the University of Pittsburgh School of Medicine. This was followed by an internship and residency in neurology at University Hospital, Ann Arbor, Michigan. He spent several years at the National Institutes of Health engaged in electron microscopy and gerontology research, has practiced clinical neurology, and, in 1983, received an M.S. degree in computer science at the University of California, San Diego. Dr. Malkoff is a member of the American Academy of Neurology, American Association for Artificial Intelligence, Institute of Electrical and Electronic Engineers, Association for Computing Machinery, Society for Neuroscience, American Society of Naval Engineers, American Association for Medical Systems and Informatics, and American Medical EEG Associates. He is certified by the American Boards of Psychiatry and Neurology, and has taught courses in neuroanatomy, clinical neurology, and computer science. At the Navy Personnel Research and Development Center, Dr. Malkoff previously worked in the area of magnetoencephalography, and is currently in charge of the Real-Time Machine Learning Laboratory where his basic interests are in the application of artificial intelligence and learning expert systems to the problems of real-time fault detection and diagnosis in complex systems.

Trend Analysis for Real-Time Stochastic Problems[1]

Donald B. Malkoff

Independent Exploratory Development

The successful performance of real-time sensor-based fault detection and diagnosis in large and complex systems is seldom achieved by operators. There is need of operator assistance in the form of automated decision aids, but effective aids have not been forthcoming. The lack of a workable method for handling temporal data (information about the timing of events) is seen as one of the key problems in this area. As part of the solution, a methodology has been developed that successfully makes use of temporal data to perform fault diagnosis in a subsystem of a Navy ship gas turbine engine propulsion unit. The methodology is embedded in a computer program designed to be used as a decision aid to assist the operator. It utilizes machine learning, is able to cope with uncertainty at several levels, works in real-time, and is developed to the point of possible application. The methodology is adaptable to other similar types of problems involving process control or pattern recognition. The approach is put forth as an example of how relatively simple and already existing techniques can be assembled into more powerful real-time diagnostic tools.

PROBLEM

Personnel assigned the task of controlling large, complex real-time systems have considerable difficulty in correctly diagnosing system malfunctions. This is particularly the case in systems where the operator must make decisions based upon sets of alarms triggered by sensor-data whose values have exceeded specified operational range limits. Information overload is almost always a significant part of this problem. However, in many of these systems, there is insufficient information for diagnosis regardless of the volume of data or the real-time constraints of the system: Many distinctly different causes of malfunctions result in identical sets of alarms. In such cases, operators are unable to render meaningful decisions. Examples of such systems in the military include nuclear and conventional power plants, mission control, sonar classification, and the strategic defense initiative.

OBJECTIVE

In these systems, there is often valuable additional information available that could be used to enhance the operator's diagnostic capabilities. The sequential order in which alarms are triggered is an example of readily available, easily employed, and highly beneficial information that is, unfortunately, seldom used for this purpose. Even more powerful is the information represented by the time at which these alarms are

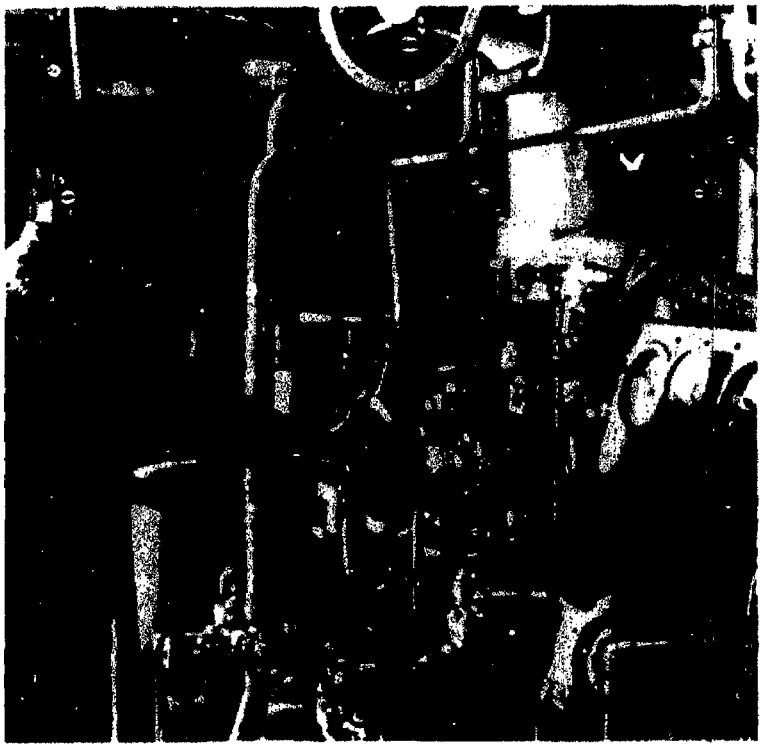
[1] Previously titled "Expert Systems for Fault Diagnosis."

triggered. The latter is not easily used to its full advantage. The facts that (1) the timing of system events is subject to considerable natural variability and that (2) the state of both the system and its environment is constantly changing present formidable obstacles to the development of a method for effectively using time for the purpose of decision making. The objective of this work is to develop and evaluate a methodology for successfully doing this, utilizing learning machine principles to overcome the problems of continual changes in the system and in the sensor inputs. Machine learning refers to the ability of a computer program to adapt automatically to changes in its inputs by modifying itself in some way so as to improve its future performance. The method must work in real-time and, therefore, in the face of incomplete information; it must be able to provide predictions as to the cause of a malfunction while the malfunction is still in the course of unfolding, as opposed to waiting until all possible untoward events have occurred.

APPROACH

A portion of the Navy surface ship gas turbine propulsion unit was selected as the testbed for this project; namely, the lubrication oil subsystem that services the reduction gears of the engine's transmission.

A computer simulator of the various malfunctions affecting this testbed was developed and tested. During the simulation, the times at which alarms are triggered were randomly varied to more realistically duplicate real-world ship situations.



Using the sequence of events and time factors to help diagnose problems with Navy machinery.

Two separate diagnostic programs were constructed:

1. A computer program was developed to utilize knowledge of the order in which alarms are triggered to diagnose the cause of malfunctions. The order was represented by a dynamic tree structure.

2. A computer program, STOCHASM, was developed to utilize knowledge of the times at which alarms are triggered to diagnose the cause of malfunction. The pattern of timing of any alarm sequence associated with a given malfunction was represented by the normal distribution of its recent values, stored in a first-in/first-out queue. Matching of

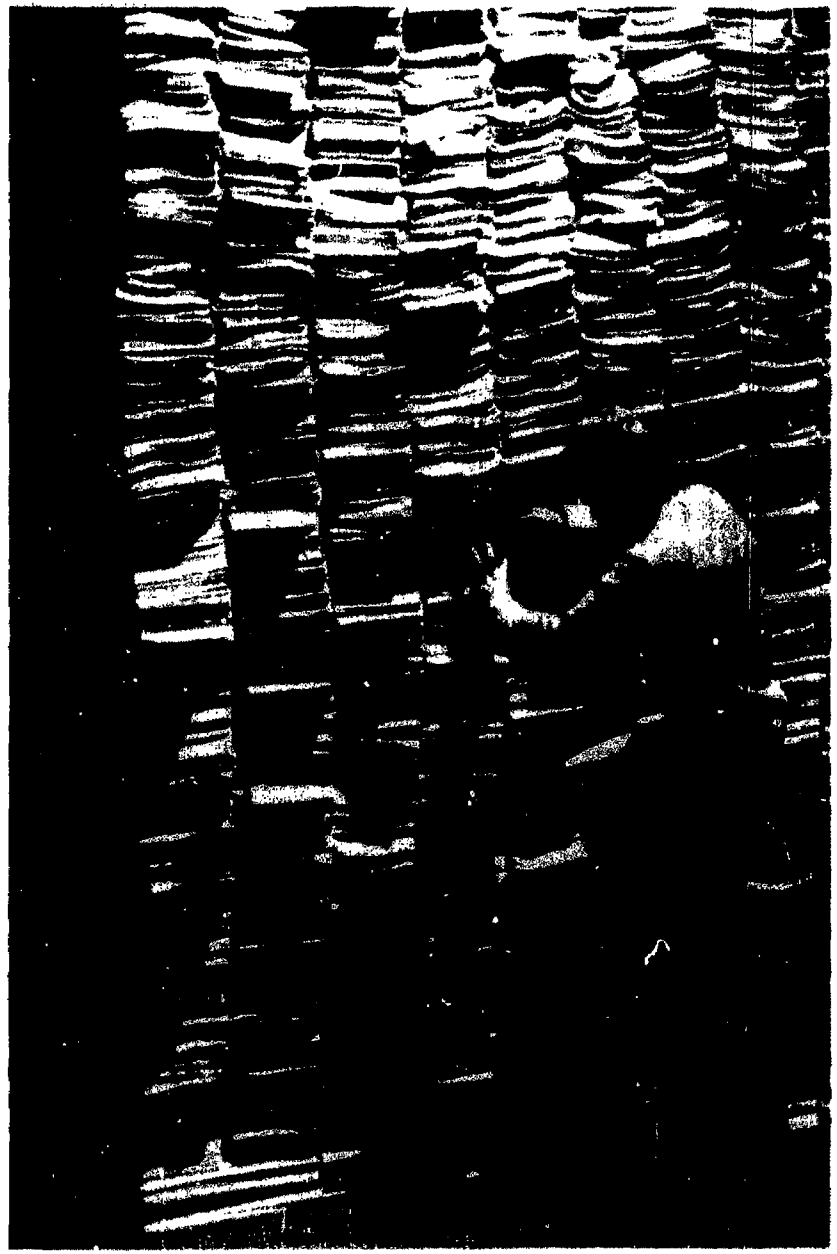
the timing characteristics of unknown malfunctions to known ones was accomplished by the use of factors based upon (1) the area under the curve of the distribution of the queued samples, delimited by the time of the latest alarm, and (2) the past incidence of the current sequence of alarms.

FINDINGS

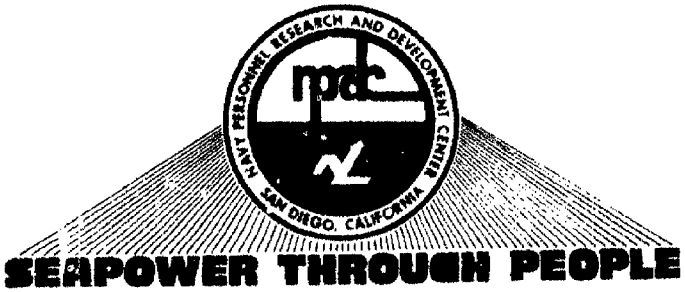
The simulator was able to duplicate the various problems inherent in the reduction gear lubrication oil subsystem. Use of knowledge of the sequence of alarms was found to significantly improve the ability to correctly diagnose the cause of a malfunction as it is unfolding. The methodology employed for handling knowledge of the timing of alarms worked effectively; it proved superior to utilization of only sequential information, worked well in real-time, and its learning capabilities successfully allowed it to automatically adapt to changes in the system and the environment. The degree of performance improvement was greatest early in the course of malfunctions, where it was most useful. Possible deficiencies in the design of the actual testbed were revealed by the program. Recommendations have been made as to future requirements for fault handling of these systems.

CONCLUSIONS

The methodology for handling temporal data to enhance real-time fault detection and diagnosis is quite effective. While the program is not yet a comprehensive one, it could provide the operator with a powerful tool to assist him in deciding the cause of malfunctions as symptoms are unfolding. The methodology is generalizable to other problem domains that involve sensor-based diagnosis in real-time under conditions of uncertainty.



Finding the needle in the haystack--analogous to determining the cause of a malfunction in complex machine systems.



William E. Montague

William Montague is Senior Scientist in the Training Laboratory at the Navy Personnel Research and Development Center (NAVPERSRANDCEN). For several years he directed projects developing improvements of instructional design methods, and incorporating computers for training. Much of this work has involved developing prescriptive methods based on research knowledge for instructional design, and ways of improving internal quality assurance for the instructional program development process. Primary efforts were the development of the Instructional Quality Inventory, a Navy test development manual, and test/evaluation of the use of microcomputer systems for instruction and instructional program development. He was trained as an experimental psychologist at the University of Virginia, did research in human factors for the Naval Electronics Laboratory, taught psychology and educational psychology at the University of Illinois, and moved to NAVPERS-RANDCEN as a project leader in 1972. He is an active member of several professional associations including American Educational Research Association, Cognitive Science Society, Psychonomic Society, American Psychological Association, Human Factors Society, and the Military Testing Association. He is past president and a director of the local chapter of the Human Factors Society. He has authored or co-authored over 100 professional and technical papers, and has co-edited three books concerned with instructional psychology.

Changes in Cognitive Structures Training

William E. Montague

Independent Exploratory Development

Two studies examined the impact of training on naive prior knowledge in the domain of basic electricity using subjects from the Navy Basic Electricity/Electronics (BE/E) course. Preliminary analyses of empirical results indicate that: (1) after subjects had successfully completed the self-paced BE/E course, their naive and incorrected knowledge persisted and may, in fact, have caused the errors observed in their reasoning, and (2) subjects who had successfully completed either a self-paced or standard classroom course still had considerable misunderstanding of circuit functioning regardless of the delivery system.

PROBLEM/BACKGROUND

Knowledge about the processes and progress of changes in cognitive structures during learning is important for making practical decisions about teaching. Methods for studying these changes are lacking and as such, research is rare. Most research has been concerned with how complex cognitive tasks are performed and the knowledge underlying that performance rather than with the development of knowledge and s[illegible] The consensus from these analyses is that expertise depends in large part on learning domain-specific structural relations between objects in the domain.

Prior work has shown that appropriate schemata can produce accurate performance, but that flaws in these mental structures cause errors in task performance; students who had completed a Navy course by successfully passing objective tests, had flawed mental structures. Some of these flaws seem to be due to the persistence of naive prior knowledge that is not eliminated by the instruction. Therefore, students were examined for their prior knowledge when they begin learning a domain and how this knowledge changes or fails to change with instruction. Since research examining mental structures is relatively new, a standard methodology does not exist. It was necessary to experiment with techniques for documenting and observing a person's knowledge structures to determine how they change.

The domain of the current effort was a course in basic electricity taught to about 25,000 Navy trainees each year. Students must pass objective tests that ostensibly examine their understanding of simple direct current and alternating current circuits, and their skill in calculating solutions to problems. The current course materials exist in two formats: one for self-paced delivery and the other for group-based, fixed-time period

delivery. Therefore, prescriptive recommendations about training procedures that derive from the research results may be used fairly directly when course materials are revised.

APPROACH

Two empirical studies were undertaken during FY86. In the first, means were developed for examining student knowledge structures and determining whether and how they change with training. Paper and pencil tests of semantic knowledge related to terminology used in electricity, direct knowledge about electricity, and qualitative understanding of relevant mathematics were devised. About 50 trainees received these tests before they started the Navy Basic Electricity/Electronics (BE/E) course, and again after they had completed the course successfully.

In the second study, about 300 trainees were assigned to one of two types of training, which consisted of self-paced or traditional classroom instruction (25 people per class). The objective was to determine whether different training methods resulted in performance differences on tests of student knowledge of circuit functioning. This effort was part of a training methods evaluation requested by the Chief of Naval Education and Training.

Computer programs were developed to deliver versions of the diagnostic tests on microcomputers (IBM-PC, Zenith 150). These programs will be tested in the future.

FINDINGS

The data have not been completely analyzed at this time. However, preliminary results are available. Empirical data from the first study shows that although there are changes in training knowledge structures in training, naive and incorrect knowledge persists and is presumably the cause of errors in reasoning. The results indicate that the course materials may not be adequate to teach the understanding of circuit functioning that trainees need as the basis for learning in later courses.

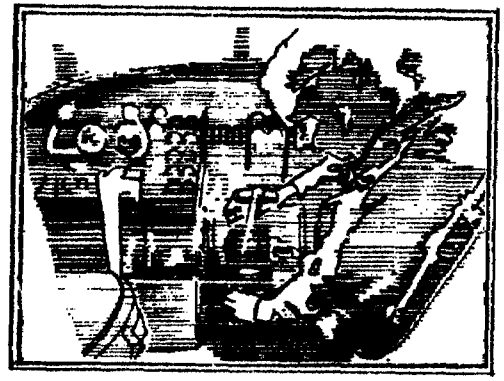
Sailors are tested on BE/E training to improve course procedures and ultimately to improve maintenance of fleet equipment.

Comparison between trainees in the second study revealed no differences attributable to the method of delivery of the materials. Students in both conditions had considerable misunderstanding of circuit functioning. Both groups made extensive errors, indicating poorly integrated cognitive knowledge structures. Less than 10 percent of the students made no errors on the diagnostic tests.

CONCLUSIONS/RECOMMENDATIONS

The preliminary conclusion is that course content probably requires considerable changes to improve student knowledge.

Methods have been developed to examine changes in semantic concepts related to electricity, qualitative understanding of mathematics, and the specific knowledge about electricity. These methods hold promise not only as techniques to assess change during training, but for helping to adapt training to individual student's knowledge. Research designed to analyze the error patterns in student responses to circuit problems has revealed the extent of student misunderstanding of the current course materials. In addition, the research has provided suggestions for both practical testing and for changes in training materials that may enhance student learning. At the least, diagnostic testing of this kind should be incorporated into course procedures. Current course tests do not reveal the serious deficiencies in student learning that exist.



APPENDIX A

Independent Research Work Units for FY86 and FY87 (PE 61152N)

Work Unit		Title	Principal Investigator	Internal Code	Telephone (619) 225- or AV 933-	FY Funding ($K) 86	87
ZR000-01 042-04	.024	Models for Calibrating Multiple-choice Items	J. B. Sympson	63	6513	30	37
	.025	Approaches to Multiple Objective Assignment	T. T. Liang	61	2371	90	74
	.026	Development of a Construct Valid Instrument for Appraising the Performance of Navy Personnel	P. J. Kidder	62	2408	0	37
ZR000-01 042-06	.029	Performance on Computer-based Cognitive Tasks	P-A. Federico	51	6434	60	60
	.030	Knowledge and Process in Adult Language Competence	F. R. Chang	51	6434	50	7
ZR000-01 042-08	.031	Action Control Theory of Motivation	L. E. Atwater	01A	6122	12	0[a]
ZR000-01 042-09	.020	Analysis of Cognition in Natural Settings	E. L. Hutchins	05B	534-1134	50	0[b]
	.028	Brain Mechanisms for Human Color Vision	L. J. Trejo	41	7424	60	51
						352	266

[a]Transitioned (see Appendix C).

[b]Completed.

APPENDIX B

Independent Exploratory Development Work Units for FY86 (PE 62766N) and FY87 (PE 62396N)

Work Unit		Title	Principal Investigator	Internal Code	Telephone (619) 225- or AV 933-	FY Funding ($K) 86	87
ZF66-512	.017	Development of Graphic Design Aids	L. M. Weitzman	05B	534-1132	43.3	0[a]
	.018	Trend Analysis for Real-Time Stochastic Problems	D. B. Malkoff	41	6617	87.8	80
	.019	Changes in Cognitive Structures Training	W. E. Montague	05	6466	51.9	46
	.020	Statistical Process Control as an Enhancement to Job and Organizational Design	S. L. Dockstader	42	6935	0	32
						183	158

[a]Transitioned (see Appendix C).

APPENDIX C

Transitioned Independent Research and Independent Exploratory Development Work Units

IR/IED Title	New Title	Sponsor	Program Element	FY87 Expected Funding ($K)
Approaches to Multiple Objective Assignment	Personnel Assignment System	CNO (OP-01)	63707N	400
Development of Graphic Design Aids	AI Tools for Development of Instruction	CNO (OP-01)	63720N	50
Action Control Theory of Motivation	Future Technologies For Manpower and Personnel Systems	ONT	62763N	20
Knowledge and Process in Adult Language Competence	Improving Job Related Literacy	ONT	62763N	100

APPENDIX D

Publications

Independent Research

Atwater, L. E., & Crawford, A. M. (1986). Leadership preparedness in newly commissioned naval officers. Proceedings of the 27th Annual Conference of the Military Testing Association, San Diego.

Atwater, L. E., Penn, R., & Trumbo, L. (1986). Leadership training at Surface Warfare Officer School. Proceedings of the Tenth Psychology in the DoD Symposium, Colorado Springs, CO.

Chang, F. R. (in press). Writing text for many readers: A role for technology. In D. Wagner (Ed.), The future of literacy in a changing world: Syntheses from the industrialized and developing nations. Elmsford, NY: Pergamon Press.

Chang, F. R. (1986, in press). Basic skills: Research with bilingual Navy recruits. In T. G. Sticht, F. G. Chang, & S. Wood (Eds.), Cognitive science and human resources management. Greenwich, CT: JAI Press.

Federico, P-A. (1986). Crystallized and fluid intelligence in a "new" instructional situation. Contemporary Educational Psychology, 11, 33-53.

Hutchins, E. (April 1986). Mediation and automatization. Quarterly Newsletter of the Laboratory of Comparative Human Cognition, 8(2).

Malkoff, D. B. (1986). Expert systems for fault diagnosis: Stochastic processes (NPRDC Spec. Rep. 86-1). San Diego: Navy Personnel Research and Development Center.

Montague, W. E. (1986). Computer-based systems for Navy training. Naval Institute Proceedings.

Montague, W. E., & Parchman, S. (1986). Procedures and problems in the Navy's BE/E course. Proceedings of the 27th Annual Conference on the Military Testing Association Conference, San Diego.

Montague, W. E., & Parchman, S. (1986). Instructional procedures and learning problems in the Navy's BE/E course. Proceedings of the 27th Annual Conference of the Military Testing Association, San Diego.

Morrison, R. F., & Hock, R. R. (1986). Career building: learning from cumulative work experience. In D. T. Hall and Associates (Ed.), Career development in organizations, pp. 236-273. San Francisco, CA: Jossey Bros.

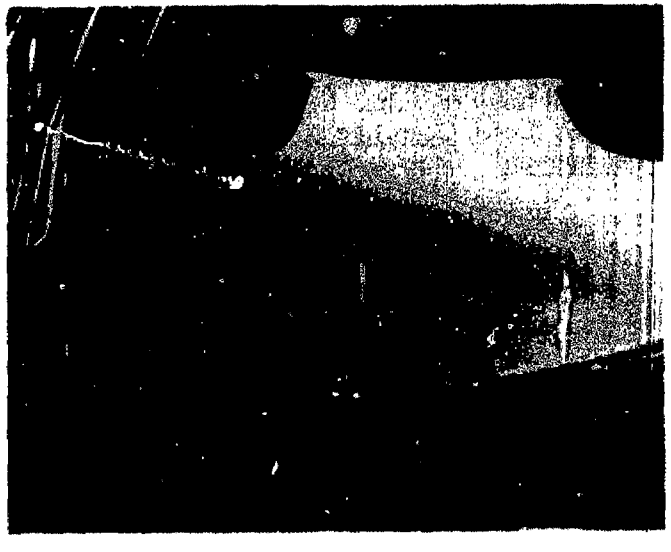
Nebeker, D. M. (1986). Effects of reward magnitude and difficulty of performance standards on individual productivity. Proceedings of the 27th Annual Conference of the Military Testing Association Conference, San Diego.

Sympson, J. B. (1986). Models for Calibrating Multiple-choice Items. In R. F. Dillon (Ed.), Advances in Testing and Training (Tech. Rep. 86-9), pp. 36-49. Carbondale, IL: Department of Psychology.

Independent Exploratory Development

Liang, T., & Lee, S. (1985). A system approach to integrate manpower planning and operations. Socio-Economic Planning Sciences, 19(6), 371-377.

Greitzer, F. L., Hershman, R. L., & Kaiwi, J. (1985). Intelligent interfaces for C^2 operability. Proceedings of the IEEE International Conference on Systems, Man, and Cybernetics.



APPENDIX E

Presentations

Independent Research

Chen, M. M., & Liang, T. T. (1986). A policy simulation model with embedded optimization and feedback system. Presentation at the Joint National Meeting of the Institute of Management Sciences and Operations Research society of America, Los Angeles.

Hutchins, E. (September 1986). Metaphors for interface design. Invited paper for the NATO Conference on the Structure of Multimodel Dialogues Including Voice, Venaco, Corsica, France.

Liang, T. T., & Buclatin, B. B. (1986). A model to improve decision making on skill training for the Navy's enlisted personnel. Presentation at the Joint National Meeting of the Institute of Management Sciences and Operations Research society of America, Los Angeles.

Sympson, J. B. (April 1986). Some item response functions obtained in polychotomous item anaiysis. Talk given at the Office of Naval Research Conference on Model-based Psychological Measurement, Gatlinburg, TN.

Sympson, J. B. (August 1986). Extracting information from wrong answers in computerized adaptive testing. Paper presented at the annual meeting of the American Psychological Association, Washington, DC.

Sympson, J. B. (September 1986). Extracting information from wrong answers in computerized adaptive testing. Invited talk at Educational Testing Service, Princeton, NJ.

Independent Exploratory Development

Chang, F. R., Gross, M. K., & Kastan, T. E. (May 1986). A lexical analyzer and parser for instructional text. Paper presented at the 4th Annual Conference on Computers and Writing, University of Pittsburgh, Pittsburgh, PA.

Greitzer, F. L., Kaiwi, J. L., & Hershman, R. L. (1985). Operator-computer interaction with intelligent systems. Third Annual Workshop on Command Decision Aiding Technology for Military Command and Control.

DISTRIBUTION LIST

Assistant Secretary of Defense (Manpower, Reserve Affairs and Logistics
Deputy Under Secretary of Defense for Research and Engineering (Research and Advanced Technology)
Military Assistant for Training and Personnel Technology (OSUDS) (R&AT)
Defense Technical Information Center (DDAC) (2)
Assistant Secretary of the Navy (Manpower & Reserve Affairs) (OASN) (M&RA) (RE&S)
Deputy Assistant Secretary of the Navy (Manpower)
Chief of Naval Operations (OP-01B7) (2), (OP-987H)
Commander, Space and Naval Warfare System Command (SPAWAR 05), (SPAWAR 05T), (SPAWAR 51E1)
Commander, Naval Civilian Personnel Command
Chief, Bureau of Medicine and Surgery (MED 25)
Commanding Officer, Naval Aerospace Medical Research Laboratory
Chief of Naval Research (Code 114) (15), (ONT 0722), (OCNR 30T)
Office of Naval Research, Detachment Pasadena
Officer of Naval Research, London
Chief of Naval Education and Training (Code OOA)
Commander, Naval Air Development Center (Code 01)
Officer in Charge, White Oak Laboratory, Naval Surface Weapons Center (U-22) (Code R04), Silver Spring, MD (2)
Commander, Naval Ocean Systems Center (Code 013) (2)
Commander, Naval Weapons Center (Code 38) (2)
Commander, David W. Taylor Naval Ship R&D Center (012) (2)
Commanding Officer, Naval Training Systems Command (Code 7A)
Commanding Officer, Naval Coastal Systems Center (Code 40)
Commanding Officer, Naval Underwater Systems Center (Code 10) (2)
Commander, Naval Air Systems Command (Code 03) (4)
Commander, Naval Sea Systems Command (003) (5)
Commandant of the Marine Corps (MPI-20) (Code RD)
Commander, Army Research Institute for the Behavioral and Social Sciences, Alexandria, VA (PERI-POT-I), (PERI-SZ)
Program Manager, Life Sciences Directorate (AFOSR/NL), Bolling AFB, DC
Commander, Air Force Human Resources Laboratory (AFHRL/MO Manpower and Personnel Division, Scientific and Technical Information Officer, Brooks AFB, TX
Commander, Air Force Human Resources Laboratory (AFHRL/LR-TDC), Wright-Patterson AFB, OH
Commander, Air Force Human Resources Laboratory (AFHRL/ID), Lowry AFB, CO
Superintendent, Naval Postgraduate School
Director of Research, U.S. Naval Academy
Executive Director, U.S. Naval Institute
Institute for Defense Analyses, Science and Technology Division
Center for Naval Analyses

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE

AD-A176 944

REPORT DOCUMENTATION PAGE

1a. REPORT SECURITY CLASSIFICATION UNCLASSIFIED	1b. RESTRICTIVE MARKINGS
2a. SECURITY CLASSIFICATION AUTHORITY	3. DISTRIBUTION/AVAILABILITY OF REPORT Approved for public release; distribution unlimited.
2b. DECLASSIFICATION/DOWNGRADING SCHEDULE	
4. PERFORMING ORGANIZATION REPORT NUMBER(S) NPRDC SR 87-1	5. MONITORING ORGANIZATION REPORT NUMBER(S)
6a. NAME OF PERFORMING ORGANIZATION Navy Personnel Research and Development Center 6b. OFFICE SYMBOL (If applicable) Code 01A	7a. NAME OF MONITORING ORGANIZATION
6c. ADDRESS (City, State, and ZIP Code) San Diego, CA 92152-6800	7b. ADDRESS (City, State, and ZIP Code)
8a. NAME OF FUNDING/SPONSORING ORGANIZATION Chief of Naval Research 8b. OFFICE SYMBOL (If applicable)	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER
8c. ADDRESS (City, State, and ZIP Code) 800 N. Quincy Street Arlington, VA 22217-5000	10. SOURCE OF FUNDING NUMBERS PROGRAM ELEMENT NO. 61152N 62396N TASK NO. WORK UNIT ACCESSION NO.

11. TITLE (Include Security Classification)

FY 86 INDEPENDENT RESEARCH/INDEPENDENT EXPLORATORY DEVELOPMENT

12. PERSONAL AUTHOR(S)
Crawford, Alice M., Metcalfe, Margen H. (Editors)

13a. TYPE OF REPORT	13b. TIME COVERED	14. DATE OF REPORT (Year, Month, Day)	15. PAGE COUNT
Annual	FROM Oct 85 TO Sep 86	January 1987	62

16. SUPPLEMENTARY NOTATION

17. COSATI CODES		18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
FIELD	GROUP	Independent research (IR), independent exploratory development (IED)
05	05, 08, 09	

19. ABSTRACT (Continue on reverse if necessary and identify by block number)

This report provides summaries of FY86 Independent Research (IR) and Independent Exploratory Development (IED) efforts and the IR/IED funding profile, and presentations and publications based on IR/IED efforts.

The following IR efforts are described: models for calibrating multiple-choice items, enhancing approaches to multiple objective assignment, performance on computer-based cognitive tasks, knowledge and process in adult language competence, analysis of cognition in natural settings, and brain mechanisms for human color vision. The following IED efforts are described: development of graphic design aids, trend analysis for real-time stochastic problems, and changes in cognitive structures training.

20. DISTRIBUTION/AVAILABILITY OF ABSTRACT	21. ABSTRACT SECURITY CLASSIFICATION
☒ UNCLASSIFIED/UNLIMITED ☐ SAME AS RPT. ☐ DTIC USERS	UNCLASSIFIED
22a. NAME OF RESPONSIBLE INDIVIDUAL Alice M. Crawford	22b. TELEPHONE (Include Area Code) (619) 225-6122 22c. OFFICE SYMBOL Code 01A

DD FORM 1473, 84 MAR 83 APR edition may be used until exhausted. All other editions are obsolete.

SECURITY CLASSIFICATION OF THIS PAGE

UNCLASSIFIED